

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 08:10:45 ; Search time 8930 Seconds  
(without alignments)  
7785.779 Million cell updates/sec

Title: US-10-071-879-9

Perfect score: 1908  
Sequence: 1 atggccccgcgcctcctgc.....tggagaagagtcgtcgtga 1908

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1523.8	79.9	3194	3 AK081560	Mus muscu
2	1423	74.6	1684	3 CR615886	full-length
3	929.6	48.7	1091	4 BM552976	AGENCOURT
4	770.8	40.4	1012	1 AL528686	AL528686
5	769	40.3	867	5 BU542726	AGENCOURT
6	745.8	39.1	1102	5 BQ063824	AGENCOURT
7	737.4	38.6	953	5 BQ538343	AGENCOURT
8	714.6	37.5	1085	1 AL563226	AL563226
9	713.8	37.4	996	4 BM915392	AGENCOURT
10	706	37.0	901	5 BQ886962	AGENCOURT
11	697.4	36.6	708	4 BQ046297	AGENCOURT
12	680.6	35.7	1086	5 BQ062401	AGENCOURT
13	679	35.6	703	4 BM784343	K-EST0062
14	670.8	35.2	674	4 BM722122	UT-E-EOO-
15	655.6	34.4	915	6 CD358944	AGENCOURT
16	653	34.2	689	4 BG328825	602427427
17	646.6	33.9	1145	4 BM811011	AGENCOURT
18	633.4	33.2	635	4 BM742719	K-EST0015
19	629.6	33.0	751	2 BE879258	K-EST0014
20	618.4	32.4	620	4 BM742162	K-EST0015
21	616.4	32.3	620	4 BM742315	K-EST0015
22	615.4	32.3	1697	9 BM744495	K-EST0018
23	615.4	32.3	1697	9 AY411504	Homo sapi
24	607.4	31.8	609	4 BM742747	K-EST0015

25	605.8	31.8	609	4 BM742740	K-EST0015
26	605	31.7	637	4 BM784328	K-EST0062
27	604.4	31.7	618	4 BM744466	K-EST0018
28	600	31.4	2388	9 AY416193	Homo sapi
29	597.4	31.3	1074	4 BM451726	AGENCOURT
30	594.6	31.2	972	5 BQ928855	AGENCOURT
31	593.4	31.1	787	5 BU707544	UT-M-FRO-
32	591	31.0	602	4 BM741814	K-EST0014
33	585	30.7	2833	3 AK044375	Mus muscu
34	581	30.5	581	4 BM783370	K-EST0061
35	580.2	30.4	3204	3 AK017594	Mus muscu
36	579.4	30.4	735	5 BQ745238	UT-M-FRO-
37	577.4	30.3	1652	9 AY411505	Par trogl
38	577	30.2	589	4 BM741945	K-EST0014
39	572	30.0	572	4 BM741785	K-EST0014
40	572	30.0	582	4 BM783606	K-EST0061
41	569	29.8	569	4 BM783585	K-EST0061
42	568.8	29.8	901	6 CB181610	AGENCOURT
43	566.6	29.7	963	4 BG328220	602427519
44	559.8	29.3	563	2 BM743912	K-EST0016
45	558	29.2	570	4 BM743028	K-EST0016

## ALIGNMENTS

RESULT 1  
LOCUS AK081560  
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:CI30040J16 product:hypothetical Immunoglobulin and major histocompatibility complex domain/leucine-rich repeat/fibronectin type III domain/ Cysteine-rich flanking region, C-terminal containing protein, full insert sequence.

ACCESSION AK081560  
VERSION AK081560.1 GI:26349238  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636

REFERENCE  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 2049374  
PUBMED 11042159

REFERENCE  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagakata, S., Saeki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashinaga, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-Format sequencing pipeline with 384 multichannel sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861

REFERENCE  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)

REFERENCE	TITLE	JOURNAL	AUTHORS	COMMENT
5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
	Nature 420, 563-573 (2002)			
	(bases 1 to 3194)			
	Adachi, J., Aizawa, K., Akimura, T., Arai, A., Hashizume, W., Fukuda, S., Furuno, M., Hiramoto, K., Hiraoka, T., Hirozane, T., Hayashida, K., Hayatsu, N., Ichi, Y., Itoh, M., Kigawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komuro, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.			
	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp), URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.			
	Please visit our web site for further details.			
	URL: http://genome.gsc.riken.jp/			
	URL: http://fantom.gsc.riken.jp/			
	Location/Qualifiers			
	1. 3194			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/strain="C57BL/6J"			
	/db_xref="PANTOM:DB:0130040J16"			
	/db_xref="taxon:10090"			
	/clone="C130040J16"			
	/tissue_type="head"			
	/clone_lib="RIKEN full-length enriched mouse cDNA library"			
	/dev_stage="16 days embryo"			
	999. 2909			
	/note="unnamed protein product, hypothetical Immunoglobulin and major histocompatibility complex domain/leucine-rich repeat/fibronectin type III domain/ cysteine-rich flanking region, C-terminal containing protein (InterPro IPR003006, evidence: InterPro)			
	putative"			
	/codon_start=1			
	/protein_id="BAC38259.1"			
	/db_xref="GI:26349239"			
	/translation="MAAPPLLLILASGAACPPLPCVCONLSISLTCARHGLLPVPEVNDRTVELRLADNFITQALGPPDFRNMTGLVDLTLSRNAIRIGARSGDLSLSLMDLGNRLVELGSSSLRGPVNLGHILISGNQRLIAGAPDFDLSLEDVDVYNNIRIQYARTAEAGTLESPAVGVTEVTATSGVSLGADPADVVMFOIQVNSBDETLIYAPASHHFLKHLVPGADVDLCLALSPAGPSDRLATRLGCAHFTLPAITLCHGLIATVGGTITAAVGVVVAALVFTVALLVGRGANGRLPIKLSHVSQNGSTSPMPKSPPPSPPPRORSCLSDLDGTGGCYGARRLRGAMARRSHVHGLLGAGRGVGGSAERLESVV"			
	31799. 3184			
	/note="putative"			
	3194			
	/note="putative"			

[illegible]

QY	1021	ACCGCGCTGGGGAGAGCTGAGGGGGCTCAACCTGCAATCGGCACCAACCCGTGCTGAGGCC	1089
Db	2019	ACGGCGCTGGAGATGACAGAGACCTATATCTTGCAATGGCCACCAACCTGCTGTGAGGCC	2078
QY	1081	AACGCGCGAGTAGAACTCGGGGTGCTGGCCTTGCCCACTGGTGGGAACAGCAGTCCGAG	1140
Db	2079	ACAGCCCGAGTGAAGCTCCGGGTACTGGCCTTGGCCCATGGTGGGAAACCCAGTCTGAG	2138
QY	1141	GGGGGCGCGCCCGGGGCGCTCGGACATGCGCCGCTCCGCTGGCAGTGTGCGGAGGTGAG	1200
Db	2139	GGGGGCGCTCTGGGCGCTTCGGAATGCTCTCTCTGTGTAAACCTGCTGGCGAAGCGAG	2198
QY	1201	GGGACGCTGAGTCTGAGCCAGCCGTGAGGTGACGAGAGTGACCGGCACCTCAAGGCTG	1260
Db	2199	GGAACTTTAGAACTGAAACAGAGGTGCAAGTGAAGAGAGTGAATGGCACTCTGGCCTG	2258
QY	1261	GTGAGCTGGGGTCCCGGGCGGCGACGCAACCAAGTGTGATGTTCCAAATCCAGTACAAC	1320
Db	2259	GTGAGCTGGGGGCTCGGGGCGGCGACGTBACCAAGTGTGAAGTTCCAAATCCAGTACAAC	2318
QY	1321	AGCAGCGAAGATGAGAACCTCATCTAACCGAATTGTCACAGGCGTCCAGCACCACTTCCGTG	1380
Db	2319	AGCAGCGAAGATGAGAACCTCATCTAACCGAATTGTAACAGCCTCCAGCACCACTTCCGTG	2378
QY	1381	CTGAAGCACTTCGTCCCGCGCTGACTATGACCTCTGCTGCTGGCCTTGTCAACCGGCC	1440
Db	2379	CTGAAGCACTCTGTCTCCCGGTGTGACTACAGACCTTGCCTGTGGGCCCTGTACACTGCT	2438
QY	1441	GCTGGGCGCTGTGACCTCAAGGCCACCAAGAGGTGCTGGGCTGGGCGCAATTTTCCACAGTG	1500
Db	2439	GCTGGGCGCTTCGACCTCAAGGCCCACTAGACTGCTGGGCTGGGCCCACTTCTCTACAGTA	2498
QY	1501	CCGGCCTTCGCCCTCTGTGCGACGCGCTTGAGGCGCCACAGCTGTGGGCGGAGACCTGACCTG	1560
Db	2499	CCAGCGACCCCGCTGTGCGACATGACCTTACAGGCCCAATGTGCTGGGCGGAGACCTTAACCTG	2558
QY	1561	GCCGTGGGGGGTGTGTCTGTGTGGCTGCTTACTGGTCTTCACTGTGGCTTGTGTGTTGG	1620
Db	2559	GCAATGGGGGGCGTCTGTAGTGGCTGCCCTTACTGGTCTTCACTGTGGCTTGTGTGTTGG	2618
QY	1621	GGCGCGGGGGGCGGAAAAATGCGCGGCTCCCGCTCAAGCTCAGCGACGTCCAGTCCAGAC	1680
Db	2619	GGCGCGGGAGGTGGGAATGAGCGCGCTCCCACTCAAACTCAGCAGTATCTCAATCCAGAC	2678
QY	1681	AATGAGGCGCCCAAGCCCAACCAAGGCCCAACCGCGCGGAGGCCCGCGCGCTCCCGGCGG	1740
Db	2679	AATGATGGCAACAGACCCCATGCTCCAAGAGCCACCAACCAAGAGCGCTCCACCCCGTCT	2738
QY	1741	CAGGCGAGCTCTCTCTGTGACCTTGGGAAATGCGC--GGGTGCTACGGTTAGTCCAGAGCGC	1797
Db	2739	CAGGCGAGCTCTCTCTGTGACCTTGGGAAACACCGAGGATGCTACGGGTATGCCAGAGCGC	2798
QY	1798	CTGGAGAGAGCTTGGGCCCGAAGGACCACTCTGTGCAATGGGGGGCTCTCGGGCAGAGG	1857
Db	2799	CTGGAGAGAGAGCTTGGGCCCGGAGGACCACTCTGTGCAATGGGGGGCTCTCGGGAGCTGG	2858
QY	1858	TGCCGCGGGGCTGAGAGAGCAGCGCCGAGGGGTCTGGAAGAGAGTGTGTGTGA	1908
Db	2859	TGTGCGGGGGGTGGGGGGGCAATGTGCGAGCGGCTTGGAAGAGAGTGTGTGTGA	2909
RESULT 2			
CRE15886			
LOCUS			
DEFINITION			
full-length cDNA clone CS00D010Y05 of Neuroblastoma Cot			
50-normalized of Homo sapiens (human).			
ACCESSION			
CRE15886			
VERSION			
CRE15886.1 GI:50496693			
KEYWORDS			
HTC, cNSLT cDNA.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Bukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK
REFERENCE	(bases 1 to 1684)	Li, W.-B., Gruber, C., Tessee, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished
REFERENCE	CONTACT : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1684)	Genoscope.	Direct Submission	Submitted (20-JUL-2004) Genoscope (E-mail : seqret@genoscope.cns.fr BP 191 91006 EVRY cedex - FRANCE)
COMMENT	Web : www.genoscope.cns.fr	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	Location/Qualifiers	1..1684
FEATURES	Source	organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"
ORIGIN	Query Match	Best Local Similarity 100.0%; Pred. No.5.4e-273; Matches 1423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	486	GGCCGGCAGCTGGGCGCCATGCTGCTGCCCTGCAACACCTTCACACTGACCATTAATTGA	545	
Qy	1	GGCCGGCAGCTGGGCGCCATGCTGCTGCCCTGCAACACCTTCACACTGACCATTAATTGA	60	
Db	546	CGCACTGCCCCAGGCGCCTTTGCCACGCTGGTCAAGCTCTCCGCTGGAAGCTCAAGCTC	605	
Qy	61	CGCACTGCCCCAGGCGCCTTTGCCACGCTGGTCAAGCTCTCCGCTGGAAGCTCAAGCTC	120	
Db	606	CAACCGGCTGGGCAAGCTGGGCTCCGAGCCGGCTTTCTCTGTTGGGGTATAGAGAGGC	665	
Qy	121	CAACCGGCTGGGCAAGCTGGGCTCCGAGCCGGCTTTCTCTGTTGGGGTATAGAGAGGC	180	
Db	666	CTCTCCGCGCCCTTGCTGCTGAGCTTTAGCGGGAACCCCTGCATGCAACTGTGAGCT	725	
Qy	181	CTCTCCGCGCCCTTGCTGCTGAGCTTTAGCGGGAACCCCTGCATGCAACTGTGAGCT	240	
Db	726	GCTGTGGCTGGGCGGCTGGGCGGCGCCGAGCAGCTTGAAACGTGGCGCTTCCCGCCGG	785	
Qy	241	GCTGTGGCTGGGCGGCTGGGCGGCGCCGAGCAGCTTGAAACGTGGCGCTTCCCGCCGG	300	
Db	786	CTGGGCGGCGCGCTACTTCTGGGGAGTGGCCGAGGGGAGATTCCTCGTGAAGCCGCCCT	845	
Qy	301	CTGGGCGGCGCGCTACTTCTGGGGAGTGGCCGAGGGGAGATTCCTCGTGAAGCCGCCCT	360	
Db	846	CATTGCCCGCACACGACGAGCGCTTGGTGGTGTGAAGGACGAGCGGCGCACAGCTCGGATG	905	
Qy	361	CATTGCCCGCACACGAGAGCGCTTGGTGGTGTGAAGGACGAGCGGCGCACAGCTCGGATG	420	
Db	906	CCGGGCGCTGGGTGACCCCGCGCTTACCATGACATGAGGATCGGTCTTACGACCGGTTGGT	965	
Qy	421	CCGGGCGCTGGGTGACCCCGCGCTTACCATGACATGAGGATCGGTCTTACGACCGGTTGGT	480	
Db	966	TGGCAACTCTCCCGGACCGCGGCTTCCCGCAACGGGACCTTGAAGATTGGGGTGAACGG	1025	
Qy	481	TGGCAACTCTCCCGGACCGCGGCTTCCCGCAACGGGACCTTGAAGATTGGGGTGAACGG	540	
Db	1026	CGCTGGGAGAGCTGGGGGCTACACTTGACATGACACCAACCTTGCTGGTGAAGGCGACAGC	1085	
Qy	541	CGCTGGGAGAGCTGGGGGCTACACTTGACATGACACCAACCTTGCTGGTGAAGGCGACAGC	600	
Db	1086	CCGAGTAGAACCTGGGGGTGCTGGCTTTGCCCATGTGGGGAACAGCATGTCGACGGAGGG	1145	

```

Db      601 CCGAGTGAAGTGGGGTCTGGGCTTGGCCCATGTTGGAGACAGCACTGCGAGGGGG 660
Qy      1146 CCGGCCCCGGGCTCTGGACATGCGCGCTCCGCTCGACATGCTGTCGAGAGGTGAAGGGAC 1205
Db      661 CCGGCCCCGGGCTCTGGACATGCGCGCTCCGCTCGACATGCTGTCGAGAGGTGAAGGGAC 720
Qy      1206 GCTGGAGTCTGAGGCGGCGGCTGAGGTGACGGAGGTGACCGGCCACTCAGGGGCTGGTGA 1265
Db      721 GCTGGAGTCTGAGGCGGCGGCTGAGGTGACGGAGGTGACCGGCCACTCAGGGGCTGGTGA 780
Qy      1266 CTTGGGGTCTCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1325
Db      781 CTTGGGGTCTCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy      1326 CGAAGATGAGACCTCATCTACCGGATTTGCTCCAGCCTCCAGCCACCACTTCTGCTGAA 1385
Db      841 CGAAGATGAGACCTCATCTACCGGATTTGCTCCAGCCTCCAGCCACCACTTCTGCTGAA 900
Qy      1386 GCACCTGCTCCCGGCGGCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1445
Db      901 GCACCTGCTCCCGGCGGCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy      1446 GCGCTGACCTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1505
Db      961 GCGCTGACCTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Qy      1506 CTGCGCCCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1565
Db      1021 CTGCGCCCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Qy      1566 GGGGGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1625
Db      1081 GGGGGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1140
Qy      1626 GGGGGCGGAAATGAGCGGCTCCCTCAAGCTGAGCGACGCTGCACTGCAAGCAATGG 1685
Db      1141 GGGGGCGGAAATGAGCGGCTCCCTCAAGCTGAGCGACGCTGCACTGCAAGCAATGG 1200
Qy      1686 AGGCGCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1745
Db      1201 AGGCGCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
Qy      1746 CAGCTGCTCTGAGACCTGAGAGTGCAGGGTGTACGGTTATGCAAGCGGCTGGAGAG 1805
Db      1261 CAGCTGCTCTGAGACCTGAGAGTGCAGGGTGTACGGTTATGCAAGCGGCTGGAGAG 1320
Qy      1806 AGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1865
Db      1321 AGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
Qy      1866 GGTGAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1908
Db      1381 GTTGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1423

```

```

RESULT 3
LOCUS   BM552976 1091 bp mRNA linear EST 20-FEB-2002
DEFINITION AGNCOURT_6572762 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466979
5', mRNA sequence.
ACCESSION BM552976
VERSION   BM552976
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1091)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  , Contact: Robert Strausberg, Ph.D.

```

Email: cgabbe-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: L10K1968 row: d column: 20  
 High quality sequence stop: 703.  
 Location/Qualifiers  
 1. 1091  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5466979"  
 /rname\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH MGC 41"  
 /note="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into BcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC library."

## ORIGIN

Query Match 48.7%; Score 929.6; DB 4; Length 1091;  
 Best Local Similarity 94.5%; Pred. No. 9.5e-175;  
 Matches 1006; Conservative 0; Mismatches 52; Indels 7; Gaps 4;

```

Qy      817 GAGGGCGAGTTCTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 876
Db      1 GAGGGCGAGTTCTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Qy      877 CTGGAAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 936
Db      61 CTGGAAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy      937 CACTGGGTGGTCTCTGACGACCGGTTGTTGCAACTCTTCCGAGCCGCGGCTTTCC 996
Db      121 CACTGGGTGGTCTCTGACGACCGGTTGTTGCAACTCTTCCGAGCCGCGGCTTTCC 180
Qy      997 AAGCGGACCTTAAGATTTGGGGTGAACCGGCGGCTGGAGGACCTGGGGCTTACACTGATC 1056
Db      181 AAGCGGACCTTAAGATTTGGGGTGAACCGGCGGCTGGAGGACCTGGGGCTTACACTGATC 240
Qy      1057 GCCACCAACCTGTGTGAGGCGGCGGCGGAGTAGAACGTGGGGGTGTGGCTTGGCC 1116
Db      241 GCCACCAACCTGTGTGAGGCGGCGGCGGAGTAGAACGTGGGGGTGTGGCTTGGCC 300
Qy      1117 CATGGTGGAAACAGCACTGCGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1176
Db      301 CATGGTGGAAACAGCACTGCGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Qy      1177 GCTCGACATGCTGCGGAGGGTGAAGGGGACGCTGGAAGTCAAGCCAGCGGTGAGTGAAG 1236
Db      361 GCTCGACATGCTGCGGAGGGTGAAGGGGACGCTGGAAGTCAAGCCAGCGGTGAGTGAAG 420
Qy      1237 GAGGTGACCGGCACTCAAGGCTGGTGAAGCTGGGGTCCGAGCGGCGGCGGCGGCGGCGG 1296
Db      421 GAGGTGACCGGCACTCAAGGCTGGTGAAGCTGGGGTCCGAGCGGCGGCGGCGGCGGCGG 480
Qy      1297 TGGATGTTCCAAATCCAGTACACAGCAGGAGATGAGACCTCATCTACCGGATTTGTC 1356
Db      481 TGGATGTTCCAAATCCAGTACACAGCAGGAGATGAGACCTCATCTACCGGATTTGTC 540
Qy      1357 CCGAGCTCCGACGACCACTTCGCTGAAGCACTGTCGCGGCGGCGGCGGCGGCGGCGGCGG 1416
Db      541 CCGAGCTCCGACGACCACTTCGCTGAAGCACTGTCGCGGCGGCGGCGGCGGCGGCGGCGG 600

```



QY 1417 TGCCTGCTGACCTTTGTCACCGCCGCTGAGCCCTCTGACCTCAAGCCACAGAGCTGCTG 1476  
 DB 601 TGCCTGCTGACCTTTGTCACCGCCGCTGAGCCCTCTGACCTCAAGCCACAGAGCTGCTG 660  
 QY 1477 GGCCTGCTGACCTTTGTCACCGCCGCTGAGCCCTCTGACCTCAAGCCACAGAGCTGCTG 1536  
 DB 661 GGCCTGCTGACCTTTGTCACCGCCGCTGAGCCCTCTGACCTCAAGCCACAGAGCTGCTG 720  
 QY 1537 GTCCTGCTGACCGCCCTGACCGCTGAGCCCTCTGAGCCCTCTGAGCCCTCTGAGCTGCTG 1596  
 DB 721 GTCCTGCTGACCGCCCTGACCGCTGAGCCCTCTGAGCCCTCTGAGCCCTCTGAGCTGCTG 780  
 QY 1597 TTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1656  
 DB 781 TTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 QY 1657 CTACAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1716  
 DB 841 CTACAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 1717 CCGCGAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 1772  
 DB 901 CCGCGAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 960  
 QY 1773 CCGCGAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 1829  
 DB 961 CCGCGAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 1020  
 QY 1830 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1874  
 DB 1021 TTTTGGCATGGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065

RESULT 4  
 AL528686 1012 bp mRNA linear EST 24-MAR-2004  
 LOCUS DEFINITION AL528686 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens  
 CDNA clone CS0PD001YC05 5-PRIME, mRNA sequence.  
 AL528686  
 VERSION AL528686.3 GI:45703748  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1012)  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 13, 2001 this sequence version replaced gi:3106536.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6064.f  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?CS0PD001AB030P1&c=6064.f.  
 Location/Qualifiers  
 1..1012  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0PD001YC05"  
 /tissue\_type="NEUROBLASTOMA COT 50-NORMALIZED"  
 /issue\_type="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V

1..1012  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0PD001YC05"  
 /tissue\_type="NEUROBLASTOMA COT 50-NORMALIZED"  
 /issue\_type="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V

ORIGIN sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 40.4%; Score 770.8; DB 1; Length 1012;  
 Best Local Similarity 90.8%; Pred. No. 4e-143;  
 Matches 781; Conservative 46; Mismatches 31; Indels 2; Gaps 2;

QY 514 CACACCTCAACCTGACCATTAACCTTATTAAGACAGTGGCCGACGAGCTTGCGCCAG 573  
 DB 30 CACACCTCAACCTGACCATTAACCTTATTAAGACAGTGGCCGACGAGCTTGCGCCAG 89  
 QY 574 CTCGCTGACCTGCTCCGCTGACCTGACCTCAACCTGACCTGACCTGACCTGACCTGAC 633  
 DB 90 CTCGCTGACCTGCTCCGCTGACCTGACCTCAACCTGACCTGACCTGACCTGACCTGAC 149  
 QY 634 CCGCTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693  
 DB 150 CCGCTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 209  
 QY 694 AGCGGGAACCCCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 753  
 DB 210 AGCGGGAACCCCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 269  
 QY 754 GACGACCTGAAAAGTGCCTGCTCCGCTGCTCCGCTGCTCCGCTGCTCCGCTGCTCCGCTG 813  
 DB 270 GACGACCTGAAAAGTGCCTGCTCCGCTGCTCCGCTGCTCCGCTGCTCCGCTGCTCCGCTG 329  
 QY 814 CCGGAGGAGGAGTTCCTCTGCTGACCGCCCTCTATTTCCCGGACACAGACGCTCTTG 873  
 DB 330 CCGGAGGAGGAGTTCCTCTGCTGACCGCCCTCTATTTCCCGGACACAGACGCTCTTG 389  
 QY 874 GTGCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 933  
 DB 390 GTGCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 449  
 QY 934 ATGCACTGAGTGTGCTGCTGACGACCGGTTGTTGCAACTCTCCGAGCCGAGCTTTC 993  
 DB 450 ATGCACTGAGTGTGCTGCTGACGACCGGTTGTTGCAACTCTCCGAGCCGAGCTTTC 509  
 QY 994 CCCAAGCGGACCTTAAGATTTGGGGTGAACCGGAGCTGGGAGACGCTGGGGCTACACTGC 1053  
 DB 510 CCCAAGCGGACCTTAAGATTTGGGGTGAACCGGAGCTGGGAGACGCTGGGGCTACACTGC 569  
 QY 1054 ATGCGCACCAACCTCTGCTGCTGAGGACGACCGGAGTGAAGTGGCGGCTGAGCTTTC 1113  
 DB 570 ATGCGCACCAACCTCTGCTGCTGAGGACGACCGGAGTGAAGTGGCGGCTGAGCTTTC 629  
 QY 1114 CCCCATGTGTGGAACAGCAGTGCAG-GGGGGGGCGCCCGGAGCTTCGACATGCGCGC 1172  
 DB 630 CCCCATGTGTGGAACAGCAGTGCAG-GGGGGGGCGCCCGGAGCTTCGACATGCGCGC 689  
 QY 1173 CTCGCTGCTGACCTGCTGCGGAGGCTGAGGAGACCTGAGCTGACCGAGCTGAGCT 1232  
 DB 690 CTCGCTGCTGACCTGCTGCGGAGGCTGAGGAGACCTGAGCTGACCGAGCTGAGCT 749  
 QY 1233 GACGAGAGTGAACCGCACTCAGAGGCTGCTGAGCTGAGGCTCCGGGCGGACCGACCC 1292  
 DB 750 GACGAGAGTGAACCGCACTCAGAGGCTGCTGAGCTGAGGCTCCGGGCGGACCGACCC 808  
 QY 1293 AGTGTGATGTTCCAAATCCAGTGAACAAGACGGAAGATGAGACCTCATCTACCGGAT 1352  
 DB 809 AGTGTGATGTTCCAAATCCAGTGAACAAGACGGAAGATGAGACCTCATCTACCGGAT 868  
 QY 1353 TGTGCTGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1372  
 DB 869 TGTGCTGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 888

RESULT 5  
 BUS42726 867 bp mRNA linear EST 13-SEP-2002  
 LOCUS DEFINITION BUS42726 BUS42726 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6574646  
 AGENCOURT\_10334726 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6574646  
 5', mRNA sequence.

ACCESSION BUS42726  
 VERSION BUS42726.1 GI:22853209  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 867)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgabds-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 DNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LNCM2770 row: e column: 14  
 High quality sequence stop: 708.  
 Location/Qualifiers

FEATURES  
 source  
 1..867  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6574646"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site:1; XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 40.3%; Score 769; DB 5; Length 867;  
 Best Local Similarity 98.3%; Pred. No. 9e-143;  
 Matches 816; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
 176 TCATCCAGGCGCCCTGAGGCGCCCTGACTTCGCGCAATGAGCGGACTGAGCTGCAC 235  
 1 TCATCCAGGCGCCCTGAGGCGCCCTGACTTCGCGCAATGAGCGGACTGAGCTGCAC 60  
 236 TGTCTGCAATGCCATCACCAGCGATTGGGGCCCGGCTTTGGGGAGCTCGAGAGCTGC 295  
 61 TGTCTGCAATGCCATCACCAGCGATTGGGGCCCGGCTTTGGGGAGCTCGAGAGCTGC 120  
 296 GTTCCCTCCGACCTTGAAGGCAACAGGCTGTGGAGGCTGGGCACTCGGAGGCTCCGGGGCC 355  
 121 GTTCCCTCCGACCTTGAAGGCAACAGGCTGTGGAGGCTGGGCACTCGGAGGCTCCGGGGCC 180  
 356 CCGTCATCTGACAGCCTCATCTCTCAGCGGCAACAGGCTGTGGAGGCTCGGATCGCGCGGAG 415  
 181 CCGTCATCTGACAGCCTCATCTCTCAGCGGCAACAGGCTGTGGAGGCTCGGATCGCGCGGAG 240  
 416 CTTTCGACGACTTCTGAGAGAGCTGAGAGCCTGACCTGTCTCAACAACACTTCGGGC 475  
 241 CTTTCGACGACTTCTGAGAGAGCTGAGAGCCTGACCTGTCTCTCAACAACACTTCGGGC 300  
 476 AGGAGCCCTGGGCGGAGATGGGGCCATGCTGCTGCTGACACCTCAACTGAGACATA 535  
 301 AGGAGCCCTGGGCGGAGATGGGGCCATGCTGCTGCTGACACCTCAACTGAGACATA 360  
 536 ACCATTATGAGCACTGCGCCAGAGGCGCTTGGCCAGCTGTGACAGCTTCCCGGCTGG 595  
 361 ACCATTATGAGCACTGCGCCAGAGGCGCTTGGCCAGCTGTGACAGCTTCCCGGCTGG 420

QY 596 ACCTCACCTCAACCGGCTGGGCAAGCTGGCTCGGACCCGCTTTCCTGTGGGCGTG 655  
 DB 421 ACCTCACCTCAACCGGCTGGGCAAGCTGGCTCGGACCCGCTTTCCTGTGGGCGTG 480  
 QY 656 ATGCAAGAGGCTTCCGCGCCCTGCTGAGCTTGAAGCGGAAACCCCTGACATGCA 715  
 DB 481 ATGCAAGAGGCTTCCGCGCCCTGCTGAGCTTGAAGCGGAAACCCCTGACATGCA 540  
 QY 716 ACTGTAGCTGCTGTGGCTGCGGAGGCTGGGCGGAGACGACCTGAAACGTCGCCCT 775  
 DB 541 ACTGTAGCTGCTGTGGCTGCGGAGGCTGGGCGGAGACGACCTGAAACGTCGCCCT 600  
 QY 776 CCGCGCCGCGGCTGGGCGGCGGCTGCTTCTGAGGAGTCCGAGGCGAGATTCCTGTG 835  
 DB 601 CCGCGCCGCGGCTGGGCGGCGGCTGCTTCTGAGGAGTCCGAGGCGAGATTCCTGTG 660  
 QY 836 AGCGGCGCTGATTTGCGCGCAACGAGGCGCTTGTGGTGAAGCGGAGCGGCA 895  
 DB 661 AGCGGCGCTGATTTGCGCGCAACGAGGCGCTTGTGGTGAAGCGGAGCGGCA 720  
 QY 896 CGCTGCGGCTGCGGCGGCTGAGACCCGCGCTCAATGACCTGGTGGT-CCTGA 953  
 DB 721 CGCTGCGGCTGCGGCGGCTGAGACCCGCGCTCAATGACCTGGTGGT-CCTGA 780  
 QY 954 GCAACCGG-TTGTGGCACT-CCTCCGAGCCCGGCTTTCCCG 996  
 DB 781 GCAACCGGTTGTGGCACTCCTCCGAGCCCGGCTTTCCCG 825

## RESULT 6

LOCUS BQ063824  
 DEFINITION 1102 bp mRNA linear EST 02-Apr-2002  
 AGENCOURT 6850285 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5925478  
 5', mRNA sequence.

ACCESSION BQ063824  
 VERSION BQ063824.1 GI:19891912  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 1102)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabds-remail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LNCM2097 row: d column: 23  
 High quality sequence stop: 689.  
 Location/Qualifiers

## FEATURES

source

1..1102  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5925478"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_99"  
 /note="Organ: lymph; Vector: pOTB7; Site:1; XhoI; Site:2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC

ORIGIN Library."

Query Match 39.1%; Score 745.8; DB 5; Length 1102;  
Best Local Similarity 94.8%; Pred. No. 3.8e-138;  
Matches 790; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 786 CTTGCGCGCGCTACTTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCCCT 845  
1 CTTGCGCGCGCTACTTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCCCT 60

QY 846 CATTGCGCGCGCAACGAGGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCG 905  
61 CATTGCGCGCGCAACGAGGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCG 120

QY 906 CCGGCGCGCTGAGTGCCTCGAGGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCG 965  
121 CCGGCGCGCTGAGTGCCTCGAGGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCG 180

QY 966 TGGCAACTCTCTCGAGGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCG 1025  
181 TGGCAACTCTCTCGAGGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGCG 240

QY 1026 CCGTGGGAGAGCTGGGAGGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 1085  
241 CCGTGGGAGAGCTGGGAGGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 300

QY 1086 CCGAGTAGAAGCTGGGAGGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 1145  
301 CCGAGTAGAAGCTGGGAGGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 360

QY 1146 CCGCGCGCGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGGAGT 1205  
361 CCGCGCGCGCGCTCTGCGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGGAGT 420

QY 1206 GCTGGAGTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 1265  
421 GCTGGAGTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 480

QY 1266 CTTGGGAGTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 1325  
481 CTTGGGAGTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAG 540

QY 1326 CCGAAGTAGAAGCTCTGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 1385  
541 CCGAAGTAGAAGCTCTGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 600

QY 1386 GAAAGTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 1445  
601 GAAAGTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 660

QY 1446 GCGCTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 1504  
661 GCGCTCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 720

QY 1505 CCTGCGCGCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 1563  
721 CCTGCGCGCTGAGGCGCGCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 780

QY 1564 GTGGGAGGAGTCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 1616  
781 GTGGGAGGAGTCTGAGGAGTGCCTCGAGGCGAGTCTCTGTAGCGCGCGAGTCTCTGTAG 833

RESULT 7  
BUS38343 953 bp mRNA linear EST 13-SEP-2002  
LOCUS AGNCOUNT\_10181110 NIH\_MGC\_107 Homo sapiens cDNA clone  
DEFINITION IMAGE:5568883 5', mRNA sequence.  
ACCESSION BUS38343  
VERSION BUS38343.1 GI:22848784  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 953)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaab@remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LNCM2755 row: e column: 11  
High quality sequence stop: 465.  
Location/Qualifiers  
1..953  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5568883"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_107"  
/note="Organ: breast; Vector: pOTB7; Site\_1: EcoRI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(9). Library constructed by  
ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 38.6%; Score 737.4; DB 5; Length 953;  
Best Local Similarity 92.9%; Pred. No. 1.8e-136;  
Matches 796; Conservative 0; Mismatches 56; Indels 5; Gaps 2;

QY 99 CTGTGCGCGCGAGGCTGCTGTTGTGCGCGCGCAAGTGGACCGCGAGTGGAGCT 158  
1 CTGTGCGCGCGAGGCTGCTGTTGTGCGCGCGCAAGTGGAGCTGAGCT 60

QY 159 GCGGCTGGTGAACCTTATCCAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 218  
61 GCGGCTGGTGAACCTTATCCAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 219 ACTGTGAGACTGAGACTGCTGCGAATGCAATCCGCAATTTGGAGGCGCGCGCGCGCGCG 278  
121 ACTGTGAGACTGAGACTGCTGCGAATGCAATCCGCAATTTGGAGGCGCGCGCGCGCGCG 180

QY 279 GGAATCTGAGAGCTTGGCTTCCCTTCACTTGAAGCGCAACAGGCTGGAGCTGGGAC 338  
181 GGAATCTGAGAGCTTGGCTTCCCTTCACTTGAAGCGCAACAGGCTGGAGCTGGGAC 240

QY 339 CGGAGCGCTCGGCG 398  
241 CGGAGCGCTCGGCG 300

QY 399 CCGCATCGGCG 458  
301 CCGCATCGGCG 360

QY 459 CTACAACAACCTTCGCGAGGCTGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 518  
361 CTACAACAACCTTCGCGAGGCTGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 420

QY 519 CCTGAACCTGAGCACTTATTTAGAGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 578  
421 CCTGAACCTGAGCACTTATTTAGAGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480





Db	1	TAGAACTGCGGGTCTGAGCCTTGGCCCATAGGTGGAGAACAGCAGTGGCCGAGGGGCGCGCC	60
Qy	1151	CCGGGACCTTGGGAATATGGCGGCTCTCCGCTCGCACTGCTGCTCCGAAGGTGAGGGAGCGCTGG	1210
Db	61	CCGGGACCTTGGGAATATGGCGGCTCTCCGCTCGCACTGCTGCTCCGAAGGTGAGGGAGCGCTGG	120
Qy	1211	AGTCTGAGCCAGCCGCTGAGGTGAGCAGAGGTGAGCCGCCACTCAGAGGGCTGTGAGTGGG	1270
Db	121	AGTCTGAGCCAGCCGCTGAGGTGAGCAGAGGTGAGCCGCCACTCAGAGGGCTGTGAGTGGG	180
Qy	1271	GTCCCCGGGCGGCCAGCGGACCCAGTGTGATGTTTCCAAATCCAGTACAAACGACCGAAG	1330
Db	181	GTCCCCGGGCGGCCAGCGGACCCAGTGTGATGTTTCCAAATCCAGTACAAACGACCGAAG	240
Qy	1331	ATGAGACCCCTCATCTACCGGATTGTCAGGCTCCAGGCCACCACTCTCTGCTGAAGCACC	1390
Db	241	ATGAGACCCCTCATCTACCGGATTGTCAGGCTCCAGGCCACCACTCTCTGCTGAAGCACC	300
Qy	1391	TGTTCCCGGCGCTGACTGACTTGTGCTCTGTGAGCCTGTGTACCCGCGCTGAGCCCT	1450
Db	301	TGTTCCCGGCGCTGACTGACTTGTGCTCTGTGAGCCTGTGTACCCGCGCTGAGCCCT	360
Qy	1451	CTGACCTCAGGGCCACCAAGCGTGTGGGCTGTGCCCATTTCTCCACGCTGCGCGCTGCG	1510
Db	361	CTGACCTCAGGGCCACCAAGCGTGTGGGCTGTGCCCATTTCTCCACGCTGCGCGCTGCG	420
Qy	1511	CCCTGTGCGCAGCCCTGTGAGGCCACAGTGTGTGGGCGGAGCCCTGACCCGTGGCGG	1570
Db	421	CCCTGTGCGCAGCCCTGTGAGGCCACAGTGTGTGGGCGGAGCCCTGACCCGTGGCGG	480
Qy	1571	GTGTGCTGT	1629
Db	481	GTGTGCTGT	540
Qy	1630	GCCGGAATAGGCGCGCTCTCCCTCAGCTCAGCAGCTCAGCTCAGCAGCAGCAGCAGCAGC	1689
Db	541	GCCGGAATAGGCGCGCTCTCCCTCAGCTCAGCAGCTCAGCTCAGCAGCAGCAGCAGCAGC	600
Qy	1690	CCCAAGCCCAACCCCAAGGCCACCC - GCGCGGAGCCCGCCCGCGCGCAGCTGAG	1748
Db	601	CCCAAGCCCAACCCCAAGGCCACCCCGCGCGGAGCCCGCCCGCGCGCGCAGCTGAG	660
Qy	1749	CTGCTCTCTGAGCTGAGCTGAGGAGATGAGC - GAGTGTCTACGTTATGAGCAGGCGCTGAGGAGAG	1807
Db	661	CTGCTCTCTGAGCTGAGCTGAGGAGATGAGC - GAGTGTCTACGTTATGAGCAGGCGCTTGTGAGGAG	720
Qy	1808	CTTGGGCGCCGACGAGACCTGTGTGATGAGGAGGAGGCTGTGAGGAGGAGGAGGAGGAGG	1867
Db	721	CTTGGGCGCCGACGAGACCTGTGTGATGAGGAGGAGGCTGTGAGGAGGAGGAGGAGGAGG	780
Qy	1868	TA 1869	
Db	781	GA 782	
RESULT 11			
LOCUS BMO46297 708 bp mRNA Linear EST 07-NOV-2001			
DEFINITION 6036260951 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452630 5',			
ACCESSION BMO46297			
KEYWORDS BMO46297.1 GI:16775564			
SOURCE EST.			
ORGANISM Homo sapiens (human)			
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE NIH-MGC http://mgi.nci.nih.gov/.			
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-remail.nih.gov			

FEATURES	SOURCE
<p>Tissue Procurement: DCD/D/DP</p> <p>CDNA Library Preparation: Ling Hong/Rubin Laboratory</p> <p>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML)</p> <p>DNA Sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: <a href="http://image.llnl.gov">http://image.llnl.gov</a></p> <p>Plate: LLCM1944 row: n column: 23</p> <p>High quality sequence stop: 704.</p> <p>Location/Qualifiers</p>	<p>1..708</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:5452630"</p> <p>/tissue_type="Carcinoma, cell line"</p> <p>/lab_host="DH10B (phage-resistant)"</p> <p>/clone_lib="NIH_MGC_40"</p> <p>/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."</p>
<p>ORIGIN</p> <p>Query Match 36.6%; Score 697.4; DB 4; Length 708;</p> <p>Best Local Similarity 99.2%; Pred. No. 1.6e-128;</p> <p>Matches 701; Conservative 0; Mismatches 6; Indels 0; Gaps 0;</p>	
<p>QY 174 CTTGATCCAGGCGCTGGGGCCCCCTGACTTCCGCAACATGACGGAGCTGTGAGACTGAC 233</p> <p>Db 2 CTTGATCCAGGCGCTGGGGCCCCCTGACTTCCGCAACATGAGCGAGCTGTGAGACTGAC 61</p>	
<p>QY 234 ACTGCTTCGCAATGACATCAACCGGATTTGGGGAGCGCGCCCTTTGGGGAGCTCGAGAGCCT 293</p> <p>Db 62 ACTGCTTCGCAATGACATCAACCGGATTTGGGGAGCGCGCCCTTTGGGGAGCTCGAGAGCCT 121</p>	
<p>QY 294 GCGTTCCCTCCACCTTGAAGGCAACAGGCTGATGAGCTGGGCACCGGAGGCTTCGAGG 353</p> <p>Db 122 GCGTTCCCTCCACCTTGAAGGCAACAGGCTGATGAGCTGGGCACCGGAGGCTTCGAGG 181</p>	
<p>QY 354 CCGCGTCAATGTGAGAGCCCTCATCTCAGGGGGAACAGAGCTGGGCGGCATTCGCGCGGG 413</p> <p>Db 182 CCGCGTCAATGTGAGAGCCCTCATCTCAGGGGGAACAGAGCTGGGCGGCATTCGCGCGGG 241</p>	
<p>QY 414 AGCCTTGACACACTTCTAGAGAGCTTGAGAGACTGTGACCTGTCTTCAACCAACCTCCG 473</p> <p>Db 242 AGCCTTGACACACTTCTAGAGAGCTTGAGAGACTGTGACCTGTCTTCAACCAACCTCCG 301</p>	
<p>QY 474 GCAGGTGCGCTTGGGCGCGCATCGGCGCATGCTGCGCTTGACACCTTCACCTGGAGCA 533</p> <p>Db 302 GCAGGTGCGCTTGGGCGCGCATCGGCGCATGCTGCGCTTGACACCTTCACCTGGAGCA 361</p>	
<p>QY 534 TAACTTATTGACGACACTGCCCGGAGGCGCTTTCGCCGACGCTCGGTACGCTTCCGCGCT 593</p> <p>Db 362 TAACTTATTGACGACACTGCCCGGAGGCGCTTTCGCCGACGCTCGGTACGCTTCCGCGCT 421</p>	
<p>QY 594 GGAACCTCACTTCACCGGCTGAGGCAAGCTGGCTTCGGAACCGGCTTTTCTCTGTGTGGG 653</p> <p>Db 422 GGAACCTCACTTCACCGGCTGAGGCAAGCTGGCTTCGGAACCGGCTTTTCTCTGTGTGGG 481</p>	
<p>QY 654 TGAGGAGAGGCGCTTCCGCGCGCGCTGAGCTTGAAGCGGGAAACCCCTGACACTG 713</p> <p>Db 482 TGAGGAGAGGCGCTTCCGCGCGCGCTGAGCTTGAAGCTTGAAGCGGGAAACCCCTGACACTG 541</p>	
<p>QY 714 CAACGTGAGCTGT 773</p> <p>Db 542 CAACGTGAGCTGT 601</p>	
<p>QY 774 CTCCCGCGCGGCTGGCGCGGCGCTACTTCTGGGCAATGCCCGAGGCGGAGTTCTCTCG 833</p>	



```

|||||
Db      602 CTCCTCCGACCGCGCTGCGCGCTACTTCTGGGAGTGGCGAGGGGAGTTCTCTG 661
Qy      834 TGAGCGCCCTGATTCGCGCCGACAGCAGCGCGCTTGGTGTCTG 880
Db      662 TGAGCTGCGCTCATTTGCGGACACAGCAGCGCTTGGTGTCTG 708

RESULT 12
LOCUS   B0062401
DEFINITION B0062401 1086 bp mRNA linear EST 02-APR-2002
5', mRNA sequence.
ACCESSION B0062401
KEYWORDS  B0062401.1 GI:19889199
SOURCE    EST.
ORGANISM  Homo sapiens (human)
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1086)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgabs-r@mail.nih.gov
           Tissue Procurement: Lou Staudt
           cDNA Library Preparation: Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNL at:
           http://image.lnl.gov
           Plate: L10M2093 row: b column: 03
           High quality sequence stop: 475.

FEATURES
Source
Location/Qualifiers
1..1086
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5923874"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 99"
/note="Organ: lymph; Vector: pORF7; Site 1: XhoI; Site 2:
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
Library."

ORIGIN
Query Match 35.7%; Score 680.6; DB 5; Length 1086;
Best Local Similarity 95.1%; Pred. No. 3.6e-125;
Matches 758; Conservative 0; Mismatches 29; Indels 10; Gaps 5;

Qy      790 GCGGCGCTACTTCTGGGAGTGGCGAGGGGAGTTCTCTGTGAGCCGCTCATTT 849
Db      1 GCGGCGCTACTTCTGGGAGTGGCGAGGGGAGTTCTCTGTGAGCCGCTCATTT 60
Qy      850 GCGGCGCACAGCAGCGCTTGGGAGTGGCGAGGGGAGTTCTCTGTGAGCCGCTCATTT 909
Db      61 GCGGCGCACAGCAGCGCTTGGGAGTGGCGAGGGGAGTTCTCTGTGAGCCGCTCATTT 120
Qy      910 GCGGCGGAGTGGCGAGCGCTTACCATGCACTGGGAGTGGCGAGCGAGTGGTGGC 969
Db      121 GCGGCGGAGTGGCGAGCGCTTACCATGCACTGGGAGTGGCGAGCGAGTGGTGGC 180
Qy      970 AATCTCTCCGAGCGCGGCTTCCCAAGGAGCTTAGAGTTGGGGTGAACCGCGCT 1029

```

```

Db      181 AACTCTCCGAGCCCGGCTTCCCAAGGAGCTTAGAGTTGGGAGTGAACCGCGCT 240
Qy      1030 GGGGAGGCTGGGGGCTACACCTGCACTGCGCAACCCCTGCTGTGAGGCGACAGCCCGA 1089
Db      241 GGGGAGGCTGGGGGCTACACCTGCACTGCGCAACCCCTGCTGTGAGGCGACAGCCCGA 300
Qy      1090 GTAGAACTGGGGGCTGGGCTTGGCCCAATGGGAGACAGCAGTGGCGGAGGGGGCGG 1149
Db      301 GTAGAACTGGGGGCTGGGCTTGGCCCAATGGGAGACAGCAGTGGCGGAGGGGGCGG 360
Qy      1150 CCCGAGCCCTCGACATGCGCCGCTCGCTCGCACTGTGCGCAGGCTGAAGGAGCGCTG 1209
Db      361 CCCGAGCCCTCGACATGCGCCGCTCGCTCGCACTGTGCGCAGGCTGAAGGAGCGCTG 420
Qy      1210 GAGTCTGAGCTCAGCCCTGCAAGTGAAGGAGTGAACCGCACTCAGGAGCTGTGAGCTGG 1269
Db      421 GAGTCTGAGCTCAGCCCTGCAAGTGAAGGAGTGAACCGCACTCAGGAGCTGTGAGCTGG 480
Qy      1270 GGTCCCGGCGGCGGACCGACCGAGTGGATGTCCTCAATCCAGTACAAAGAGCGGAA 1329
Db      481 GGTCCCGGCGGCGGACCGACCGAGTGGATGTCCTCAATCCAGTACAAAGAGCGGAA 540
Qy      1330 GATGAGACCTCTCATATCAACCGATTGTCAGACCTTCAGCACCATTCTGTGAAGCAG 1389
Db      541 GATGAGACCTCTCATATCAACCGATTGTCAGACCTTCAGCACCATTCTGTGAAGCAG 600
Qy      1390 CTGTGCTCCGCGGCGTACATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1449
Db      601 CTGTGCTCCGCGGCGTACATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy      1450 TCTGACCTGAC-GGACACAGGCTGTGGG-CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1504
Db      661 TCTGACCTGACAGGCGACAGGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy      1505 CCTGCGCCCTGTGTCACAGCCCTGCAAGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGG 1559
Db      721 CCTGCGCCCTGTGTCACAGCCCTGCAAGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 780
Qy      1560 GCGCGCTGGGGGGGTGTCG 1576
Db      781 GCGCGCTGGGGGGGTGTCG 797

RESULT 13
LOCUS   BM784343
DEFINITION K-EST0062439 S7SNUT719 Homo sapiens cDNA clone S7SNUT719-34-C06 5',
           mRNA sequence.
ACCESSION BM784343
VERSION   BM784343.1 GI:19132575
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 703)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: Yongsung@mail.krribb.re.kr
           Plate: 34 row: C column: 06
           High quality sequence stop: 703.

FEATURES
Source
Location/Qualifiers
1..703
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S7SNU719-34-C06"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-719"
/lab_host="TOP10F"
/clone_lib="S7SNU719"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
Site_2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
interact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transposition of
competent cells E. coli TOP10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

```

## ORIGIN

```

Query Match      35.6%; Score 679; DB 4; Length 703;
Best Local Similarity 99.7%; Pred. No. 7.4e-125;
Matches 701; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 372 CCTCATCTCAGAGGGAACCAAGCTGGGCGGATGCGCGGAGAGCTTCGACGACTTCCT
DB 1 CCTCATCTCAGAGGGAACCAAGCTGGGCGGATGCGCGGAGAGCTTCGACGACTTCCT
QY 432 AGAGAGCTTGAAGGAGCTTGAAGCTTCTTCTTCAACAACCTCCGAGGTGCTTGGGCGG
DB 61 AGAGAGCTTGAAGGAGCTTGAAGCTTCTTCTTCAACAACCTCCGAGGTGCTTGGGCGG
QY 492 CATGGGGGCAATGCTGCTGCTTGAACACCTTCAACCTTGAACCTTATTTAGCGACT
DB 121 CATGGGGGCAATGCTGCTGCTTGAACACCTTCAACCTTGAACCTTATTTAGCGACT
QY 552 GCGCCGAGGCGCTTGGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 181 GCGCCGAGGCGCTTGGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY 612 CTTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 241 CTTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY 672 GCGCCGAGGCGCTTGGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 301 GCGCCGAGGCGCTTGGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY 732 GCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 361 GCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY 792 GCGCGGCTACTTCTGGGAGTGGCCGAGGCGGAGTTCCTGTGAGCGCCCTCATTTGC
DB 421 GCGCGGCTACTTCTGGGAGTGGCCGAGGCGGAGTTCCTGTGAGCGCCCTCATTTGC
QY 852 CCGGCAACGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 481 CCGGCAACGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY 912 CTTGGGAGGCGGCGGCTTGAACGAGCTTGAACGAGCTTGAACGAGCTTGAACGAGCTTGAACG
DB 541 CTTGGGAGGCGGCGGCTTGAACGAGCTTGAACGAGCTTGAACGAGCTTGAACGAGCTTGAACG
QY 970 AACTCTCCGAGGCGGCGGCTTTCGCCAAGGAGCTTGAAGATTGGGAGTGAACGCGGCT 1029

```

```

DB 601 AACTCTCCGAGGCGGCGGCTTTCGCCAAGGAGCTTGAAGATTGGGAGTGAACGCGGCT 660
QY 1030 GGGAGGCGCTGGGGCTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 661 GGGAGGCGCTGGGGCTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

```

## RESULT 14

```

LOCUS BM722122 674 bp mRNA linear EST 01-MAR-2002
DEFINITION U1-E-E00-ahy-a-11-0-U1.r1 U1-E-E00 Homo sapiens cDNA clone
ACCESSION U1-E-E00-ahy-a-11-0-U1 5', mRNA sequence.
VERSION BM722122
KEYWORDS BM722122.1 GI:19042482
SOURCE EST.
ORGANISM Homo sapiens (human)

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 674)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalizaiton and subtrraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

```

```

JOURNAL MEDLINE
PUBMED 8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
315 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu

```

```

Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

```

## FEATURES

## source

```

1..674
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="U1-E-E00-ahy-a-11-0-U1"
/tissue_type="fetal eye"
/dev_stage="fetal"
/clone_lib="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
U1-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dt primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GCGGTATACC. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

## ORIGIN

```

Query Match      35.2%; Score 670.8; DB 4; Length 674;
Best Local Similarity 99.7%; Pred. No. 3.2e-123;
Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 289 AGCTGCGTTCCCTCAGCTTGAACGAGGCTGTGAGCTGGGACCGGAGGCTC 348

```

```

Db      1 AGCTGGCTTCCCTCCACCTTGAAGGCAAGGCTGTGAGCTGGGCAACCGGAGCCTC 60
QY      349 CGGAGCCCTCATCATCTGACGACCTATCTCATGCGGCAACGAGCTGGGCGCATCGCG 408
Db      61 CGGGGCCCCCTCAATCTGACGACCTCATCTCATGCGGCAACGAGCTGGGCGCATCGCG 120
QY      409 CGGAGACCTTGAAGCACTTCTGAAGAGCTGTGAGAGCTGTGCTTCTCAACAAC 468
Db      121 CGGGAGACCTTGAAGCACTTCTGAAGAGCTGTGAGAGCTGTGCTTCTCAACAAC 180
QY      469 CTCCGAGAGTGGCTGGGCGGCGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGG 528
Db      181 CTCCGAGAGTGGCTGGGCGGCGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGG 240
QY      529 GACCATAACTTATTAAGCACTGACGACCTGACGACCTGACGACCTGACGACCTGACG 588
Db      241 GACCATAACTTATTAAGCACTGACGACCTGACGACCTGACGACCTGACGACCTGACG 300
QY      589 CGCTGAGACCTCACTTCAACCGGCTGAGCGACGCTGAGCGACGCTGAGCGACGCTG 648
Db      301 CGCTGAGACCTCACTTCAACCGGCTGAGCGACGCTGAGCGACGCTGAGCGACGCTG 360
QY      649 GGGCGGATGCAAGGCGCTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 708
Db      361 GGGCGGATGCAAGGCGCTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420
QY      709 CACTGCAACTGTGTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 768
Db      421 CACTGCAACTGTGTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY      769 TGGGCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 828
Db      481 TGGGCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 540
QY      829 TCTGTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 888
Db      541 TCTGTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
QY      889 CGGAGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 948
Db      601 CGGGCGACGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY      949 CCGTGAAGACCGGTT 962
Db      661 CCGTGAAGACCGGTT 674

RESULT 15
CD358944      915 bp      mRNA      linear      EST 29-MAY-2003
LOCUS      CD358944
DEFINITION      AGENCOURT 14276726 NIH MGC 180 Homo sapiens cDNA clone
ACCESSION      IMAGE:30389020 5', mRNA sequence.
VERSION      CD358944
KEYWORDS      CD358944.1 GI:31130355
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 915)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

```

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30389020"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances )"
/clone_lib="NIH MGC 180"
/note="Organ: Testis; Vector: pCMV-Sport6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match      34.4%; Score 655.6; DB 6; Length 915;
Best Local Similarity 98.2%; Pred. No. 3.4e-120;
Matches 695; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

616 GCCACGCTGGCTCCGACCCGCTTCTCGTGGGCGTGAATGACAGGCGCTTCCCGCC 675
Db      5 GGGCGCGCTGGCTCTGACCCCGCTTCTCGTGGGCGTGAATGACAGGCGCTTCCCGCC 64
QY      676 CCCCTGTGTCTGAGCTTTAGCGGAAACCCCTGACATGCACTGTGAGCTGTGTGTGT 735
Db      65 CCCCTGTGTCTGAGCTTTAGCGGAAACCCCTGACATGCACTGTGAGCTGTGTGTGT 124
QY      736 CGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 795
Db      125 CGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 184
QY      796 CGCTACTTCTGAGGAGATGCCGAGGCGGAGTCTCTGTGAGCGGCGGCGGCGGCGGCG 855
Db      185 CGCTACTTCTGAGGAGATGCCGAGGCGGAGTCTCTGTGAGCGGCGGCGGCGGCGGCG 244
QY      856 CACAGCAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 915
Db      245 CACAGCAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 304
QY      916 GGTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 975
Db      305 GGTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 364
QY      976 TCCCGAGCCCGGCGGCTTCCCAACGAGCACTTGAAGATGTGGGTTGACCGGCGTGGGAGC 1035
Db      365 TCCCGAGCCCGGCGGCTTCCCAACGAGCACTTGAAGATGTGGGTTGACCGGCGTGGGAGC 424
QY      1036 GCTGTGGGCGGTACACCTGTGACGCAACCAACCTGTGTGTGTGTGTGTGTGTGTGT 1095
Db      425 GCTGTGGGCGGTACACCTGTGACGCAACCAACCTGTGTGTGTGTGTGTGTGTGTGT 484
QY      1096 CTGCGGCGTGTGGCTTGTGGCGCATGTGTGGGAAACAGAGTCCGAGGCGGCGGCGGCGG 1155
Db      485 CTGCGGCGTGTGGCTTGTGGCGCATGTGTGGGAAACAGAGTCCGAGGCGGCGGCGGCGG 544
QY      1156 CCTCGGACATGCGCGGCTTCCGCTGCACTGTGTGCGGA-GGGTGAAGGAGACCTGTGAGTTC 1214
Db      545 CCTCGGACATGCGCGGCTTCCGCTGCACTGTGTGCGGAAGGGGTGAAGGAGACCTGTGAGTTC 604
QY      1215 TAGAGCAAGCGGTGACAGGTGACAGAGTGAACCGGCACTTGAAGGCTGTGTGTGTGT 1274
Db      605 TAGAGCAAGCGGTGACAGGTGACAGAGTGAACCGGCACTTGAAGGCTGTGTGTGTGT 664
QY      1275 C-GGGGCGGCGAGCGGACCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1319
Db      665 CGGGGCGGCGAGCGGACCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 712

```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: ND4M461 row: n column: 05

High quality sequence start: 6

High quality sequence stop: 595.

Location/Qualifiers

1..915

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30389020"

/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances )"

/clone\_lib="NIH MGC 180"

/note="Organ: Testis; Vector: pCMV-Sport6.1; Site 1: NotI;

Site 2: EcoRV (destroyed); Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.68 kb. Library was

constructed by (Invitrogen). Note: this is a NIH\_MGC

Library."

Sat Nov 20 16:15:42 2004

us-10-071-879-9.rst

Page 14

Search completed: November 20, 2004, 14:47:35  
Job time : 8936 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 04:30:49 ; Search time 1327 seconds  
(without alignments)  
7547.773 Million cell updates/sec

Title: US-10-071-879-9

Perfect score: 1908

Sequence: 1 atggcccgccgcgtcctcgtc.....tggaagagagtggtgtga 1908

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4134886 segs, 2624710521 residues 8269772

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N\_Geneseq\_23Sep04:\*

- 1: geneseqn1808:\*
- 2: geneseqn1908:\*
- 3: geneseqn2000:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1906.4	99.9	2369	6 AAD38695	Aad38695 Human LP2
2	1906.4	99.9	2569	10 AAK96609	Aak96609 MDDT rela
3	1904.8	99.8	2364	10 ADE07450	Ade07450 Novel cod
4	1897	99.4	3201	6 ABV99420	Abv99420 Human NOV
5	1789.6	93.8	2885	10 ADI21820	Adi21820 Novel hum
6	1787.6	93.7	2563	6 ABV99422	Abv99422 Human NOV
7	1629	85.4	2451	6 AAD28124	Aad28124 Human leu
8	1584.4	83.0	1872	4 ABV99421	Abv99421 Human NOV
9	1563.4	81.9	1872	4 AAF76853	Aaf76853 Human sec
10	1379.4	72.3	1582	5 AAS69014	Aas69014 DNA encod
11	1101.8	57.7	1653	6 AAD28125	Aad28125 Human leu
12	624.8	32.7	1887	4 AAI66986	Aai66986 Human LRR
13	624.8	32.7	1887	4 AAI78215	Aai78215 Nucleotid
14	624.8	32.7	1887	4 AAI25568	Aai25568 Nucleotid
15	624.8	32.7	1887	4 ADH71653	Adh71653 Human gen
16	624.8	32.7	1887	12 ADH71651	Adh71651 Human gen
17	624.8	32.7	2549	10 ADF69163	Adf69163 Human MP5
18	624.8	32.7	2637	4 AAI66985	Aai66985 Human leu
19	624.8	32.7	3124	6 AB211543	Ab211543 Human pol
20	624.8	32.7	3245	12 AAD44061	Aad44061 Novel hum
21	621	32.5	3144	3 AAE60605	Aae60605 Human hno

22	621	32.5	3144	3 AAA40083	Aaa40083 Human bra
23	621	32.5	3144	10 ADF69162	Adf69162 Human MP5
24	619.4	32.5	2818	5 AAI58215	Aai58215 Human pol
25	619.4	32.5	2818	5 AAS44992	Aas44992 cDNA enco
26	619.4	32.5	2818	5 ADQ98422	Adq98422 DNA enco
27	619.4	32.5	2818	5 ADB48182	Adb48182 Novel hum
28	617	32.3	2316	4 AAI78204	Aai78204 Nucleotid
29	617	32.3	2869	10 AAD54310	Aad54310 Human SEC
30	617	32.3	3386	6 ABR08315	Abt08315 Human NOV
31	617	32.3	3568	6 ABR08317	Abt08317 Human NOV
32	617	32.3	4042	10 ADI21778	Adi21778 Novel hum
33	612	32.1	2855	6 ABR70010	Abk70010 cDNA enco
34	612	32.1	2855	9 ADA01367	Ada01367 Human PRO
35	612	32.1	2855	9 ADA43796	Ada43796 Human CDN
36	612	32.1	2855	9 ADA43564	Ada43564 Human CDN
37	612	32.1	2855	9 ADA01239	Ada01239 Human PRO
38	612	32.1	2855	9 ADA01123	Ada01123 Human CDN
39	612	32.1	2855	9 ADA43680	Ada43680 Human CDN
40	612	32.1	2855	9 ADA06942	Ada06942 Human PRO
41	612	32.1	2855	9 ADA08430	Ada08430 Novel hum
42	612	32.1	2855	9 ADB99723	Adb99723 Human PRO
43	612	32.1	2855	9 ADB87006	Adb87006 Human PRO
44	612	32.1	2855	9 ADB66161	Adb66161 Human CDN
45	612	32.1	2855	10 ADB99839	Adb99839 Human PRO

## ALIGNMENTS

RESULT 1	AAD38695
ID	AAD38695 standard; cDNA; 2369 BP.
XX	XX
AC	AAAD38695;
XX	XX
DT	23-SEP-2002 (first entry)
XX	XX
DB	Human LP220 secreted protein encoding cDNA.
XX	XX
KW	Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;
KW	diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KW	rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KW	reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
KW	gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
XX	XX
OS	Homo sapiens.
XX	XX
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "Human LP220 secreted protein"
FT	sig_peptide
FT	167..214
FT	/*tag= b
FT	mat_peptide
FT	215..2071
FT	/*tag= c
FT	/product= "Mature human LP220 secreted protein"
XX	XX
PN	WO200228601-A2.
XX	XX
PD	04-APR-2002.
XX	XX
PF	14-SEP-2001; 2001WO-US026026.
XX	XX
PR	28-SEP-2000; 2000US-0236088P.
XX	XX
PA	(EHLI ) LILLY & CO EHLI.
XX	XX
PI	Su EW, Wang H;
XX	XX
DR	WPI; 2002-471259/50.
DR	P-PSDB; AAE23980.
XX	XX





Db	1907	CAGCCAGCTGCTCTCTTGACCTGGAGAGATGCCGGGTGTACAGTTATGCCAGGCGCTG	1906
Qy	1801	GGAGGAGCTTGGGGCCCGACGAGCCACTCTGTGCATGGGGGGCTCTCTCGGGCAGGGTCC	186C
Db	1967	GGAGGAGCTTGGGGCCCGACGAGCCACTCTGTGCATGGGGGGCTCTCTCGGGCAGGGTCC	2026
Qy	1861	CGGGGGGTAGAGGAGCAGCGCCAGCGGCTGTGAAGAAGAGTGTGTGTGA	1908
Db	2027	CGGGGGGTAGAGGAGCAGCGCCAGCGGCTGTGAAGAAGAGTGTGTGTGA	2074
RESULT 2			
AAK99609			
AC	AAK99609	standard; DNA; 2569 BP.	
XX			
DT	10-APR-2003	(first entry)	
XX			
DE	MDPT	related human DNA SEQ ID NO 73.	
XX			
KM	Cycostatic;	antiatherosclerotic; osteoporathic; antiarteriosclerotic;	
KM	hepatotropic;	antiparasitic; antiallergic; antianemic; antiasthmatic;	
KM	antihyproid;	antiinflammatory; antihelminthic; antidiabetic; nephrotropic;	
KM	ophthalmological;	immunosuppressive; dermatological; antituber;	
KM	antirheumatic;	antiarthritic; antibacterial; virocidic; fungicide;	
KM	antiparasitic;	protozoacide; tranquilliser; vulnerary; anti-HIV;	
KM	neotropic;	neuroprotective; anticonvulsant; cerebroprotective;	
KM	neuroleptic;	molecules for disease detection and treatment; MDDT;	
KM	immunogen;	cancer; actinic keratosis; arteriosclerosis; atherosclerosis;	
KM	bursitis;	cirrhosis; hepatitis; psoriasis; AIDS; rheumatoid arthritis;	
KM	adult respiratory distress syndrome;	Addison's disease; allergy; anaemia;	
KM	asthma;	osteoporosis; autoimmune; haemolytic anaemia; scleroderma;	
KM	autoimmune thyroiditis;	Crohn's disease; atopic dermatitis;	
KM	diabetes mellitus;	Graves' disease; glomerulonephritis;	
KM	systemic lupus erythematosus;	systemic sclerosis; ulcerative colitis;	
KM	haemodialysis;	uveitis; trauma; Alzheimer's; Pick disease;	
KM	Parkinson disease;	amyotrophic lateral sclerosis; epilepsy; stroke;	
KM	Huntington's disease;	multiple sclerosis; dementia;	
KM	extrapyramidal disorder;	motor neuron disorder; central nervous system;	
KM	neuromuscular disorder;	metabolic; endocrine; toxic myopathy;	
KM	periodic paralysis;	mental disorder; human; gene; ds.	
OS	Homo sapiens.		
XX			
XX	MO200296951-AL.		
PD	05-DEC-2002.		
XX			
PF	24-MAY-2002;	2002WC-US016676.	
XX			
PR	25-MAY-2001;	2001US-0293723P.	
PR	01-JUN-2001;	2001US-0295257P.	
PR	08-JUN-2001;	2001US-0297220P.	
PR	21-JUN-2001;	2001US-0300526P.	
PR	29-JUN-2001;	2001US-0301874P.	
PR	22-FEB-2002;	2002US-0359413P.	
XX			
PA	(INCY -)	INCYTE GENOMICS INC.	
XX			
PI	Tang TY,	Yue H., Baughn MR., Duggan BM., Warren BA., Bandman O.,	
PI	Richardson TW,	Burford N., Sanjinnala B., Becha SD., Yao MG., Yang J.,	
PI	Piran UK,	Hatala AJA., Griffin JA., Swarnakar A., Elliott VS.,	
PI	Reidipon SA,	Khan FA., Lee EA., Yue H., Lu DAM., Walla NK., Thangavelu K.,	
PI	Arizawa CS,	Xu Y., Ison CH., Huang J., Ding L., Honchell CD,	
PI	Borowicz ML,	Emerling BM., Peterson DP., Lu Y., Ramkumar J., Mason PM;	
PI	Zebartadjan Y,	Azimzal Y., Stuve LL., Kamigaki LL., Barroso I., Lee S;	
XX	Kable AE;		
DR	WPI:	2003-140448/13.	
XX	P-PSDB:	AAO26256.	
DT	Novel molecules for disease detection and treatment and polynucleotide		

PT	encoding them useful for diagnosing, preventing or treating cell
PT	proliferative, autoimmune/inflammatory, neurological and developmental
PT	disorders.
XX	Claim 128; Page 255; 260pp; English.
XX	
CC	The invention relates to an isolated polypeptide chosen from molecules
CC	for disease detection and treatment (MDPT), comprising a one of 39 114-
CC	1250 residue amino acid sequences, given in the specification, or a
CC	biologically active or immunogenic fragment of the isolated polypeptide.
CC	The isolated polypeptide is useful for screening a compound for
CC	effectiveness as an agonist or antagonist of the isolated polypeptide.
CC	The isolated polypeptide is also useful as an immunogen for preparing
CC	polyclonal or monoclonal antibodies by hybridoma technology. The isolated
CC	polypeptide and its encoding polynucleotide are useful for diagnosis,
CC	treatment and prevention of cancer, including cancer, for diagnosis,
CC	atherosclerosis, burns, cirrhosis, hepatitis, psoriasis, AIDS, adult
CC	respiratory distress syndrome, Addison's disease, allergies, anaemia,
CC	asthma, atherosclerosis, osteoporosis, autoimmune haemolytic anaemia,
CC	autoimmune thyroiditis, Crohn's disease, atopic dermatitis, diabetes
CC	mellitus, Graves' disease, glomerulonephritis, rheumatoid arthritis,
CC	scleroderma, systemic lupus erythematosus, systemic sclerosis, ulcerative
CC	colitis, haemodialysis, uveitis, viral, bacterial, fungal, parasitic,
CC	protozoal, hematologic infections, trauma, Alzheimer's and Pick disease,
CC	Parkinson's disease, amyotrophic lateral sclerosis, epilepsy, stroke,
CC	Huntington's disease, multiple sclerosis, dementia, and other
CC	extrapyramidal disorder, motor neuron disorder, and other developmental
CC	disorders of the central nervous system, neuromuscular disorders,
CC	metabolic, endocrine and toxic myopathies, periodic paralysis, mental
CC	disorders including mood, anxiety and schizophrenic disorders, anaemia,
CC	renal tubular acidosis, epilepsy, hypothyroidism, glaucoma, sensorineural
CC	hearing loss and cataract. This polynucleotide sequence represents the
CC	DNA encoding a human MDPPT protein relating to the invention
XX	
XX	Sequence 2569 BP; 361 A; 928 C; 818 G; 461 T; 0 U; 1 Other;
XX	
XX	Query Match 99.9%; Score 1906.4; DB 10; Length 2569;
XX	Best Local Similarity 99.9%; Pred. No. 0;
XX	Matches 1907; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 ATGGCGCCGCGCGCTCTGCTGCTGTCGACAGTGGAGGCGCGCGCTGCGCCG
DB	367 ATGGCGCCGCGCGCTCTGCTGCTGTCGACAGTGGAGGCGCGCGCTGCGCCG
OY	61 TGGCTTGGCGGAAGCTGTCCGAGTGGCTGACAGCCCTGTGGCCACCGAGGCTGCTG
DB	427 TGGCTTGGCGGAAGCTGTCCGAGTGGCTGACAGCCCTGTGGCCACCGAGGCTGCTG
OY	121 TTGTGGCGCGCCAGTGGAGCCGGCGGACAGTGGAGCTGCGGCTGGCTGACAACTTATC
DB	487 TTGTGGCGCGCCAGTGGAGCCGGCGGACAGTGGAGCTGCGGCTGGCTGACAACTTATC
OY	181 CAGGCGCTGGGGCCCGCTGACTTCCGGAACATGACGGGACTGGTGGACCTGACATGCT
DB	547 CAGGCGCTGGGGCCCGCTGACTTCCGGAACATGACGGGACTGGTGGACCTGACATGCT
OY	241 CGCATATGCATACCCGCGATTGGGGCCCGCGCCCTTTGGGGAACCTGGAGAGCTTGCCTCC
DB	607 CGCATATGCATACCCGCGATTGGGGCCCGCGCCCTTTGGGGAACCTGGAGAGCTTGCCTCC
OY	301 CTCACCTTGGACGGACAGAGGCTGGTGGAGCTGGGACCGGGAGCGTCCCGGGGCCGCTC
DB	667 CTCACCTTGGACGGACAGAGGCTGGTGGAGCTGGGACCGGGAGCGTCCCGGGGCCGCTC
OY	361 AATTTGACGACCTCATCTTCAAGCGGACCAAGCTGGGCGCGATCGCGCGGGAGCCTTC
DB	727 AATTTGACGACCTCATCTTCAAGCGGACCAAGCTGGGCGCGATCGCGCGGGAGCCTTC
OY	421 GAGCAGCTTCTAGAGAGCCTGGAGGACCTGGAGCTGTCTTACAAACACTCTCCGCAAGTGT
DB	787 GAGCAGCTTCTAGAGAGCCTGGAGGACCTGGAGCTGTCTTACAAACACTCTCCGCAAGTGT
OY	481 CCTGGGCGGAGCATGGGCGCATATGCGCCCTGACACCTTCAACTTGAACATTAACCTT
DB	540 CCTGGGCGGAGCATGGGCGCATATGCGCCCTGACACCTTCAACTTGAACATTAACCTT

Db 847 CCTGGACCGGACATGGAGCCATGCTGCTGCAACCTCTGAACCTGACATACCTT 906  
QY 541 ATTGACGACTGCCCCCGAGGCGCTTGCCAGACCTGGTGAAGCTCTCCGCGCTGAACCTC 600  
Db 907 ATTGACGACTGCCCCCGAGGCGCTTGCCAGACCTGGTGAAGCTCTCCGCGCTGAACCTC 966  
QY 601 ACCTTCACCGCTGGGCGACAGCTGGCTCCGGAACCCGTTTTCTCTGGTGGGCGTGATGCA 660  
Db 967 ACCTTCACCGCTGGGCGACAGCTGGCTCCGGAACCCGTTTTCTCTGGTGGGCGTGATGCA 1026  
QY 661 GAGGCGCTCTCCGCGCGCTGGTGGTGAAGCTTTAGCGGGAACCCCTGCACTGCAACTGT 720  
Db 1027 GAGGCGCTCTCCGCGCGCTGGTGGTGAAGCTTTAGCGGGAACCCCTGCACTGCAACTGT 1086  
QY 721 GAGCTGCTGGTGGCGCGGCTGGCGCGGCGCGACGACTGGAAACGTGCGCTCCCG 780  
Db 1087 GAGCTGCTGGTGGCGCGGCTGGCGCGGCGCGACGACTGGAAACGTGCGCTCCCG 1146  
QY 781 CCGGCGCTGGCGCGGCGCTGACTTCTGGGCGAGTCCGAGGGCGAGTTCTCTGGAGCCG 840  
Db 1147 CCGGCGCTGGCGCGGCGCTGACTTCTGGGCGAGTCCGAGGGCGAGTTCTCTGGAGCCG 1206  
QY 841 CCCCTGATTGCGCGGCGACAGCGAGCGCTCTGGGCTGGAGGCGGACGCGCGCTG 900  
Db 1207 CCCCTGATTGCGCGGCGACAGCGAGCGCTCTGGGCTGGAGGCGGACGCGCGCTG 1266  
QY 901 CGGTGCGCGGCGCTGGGCTGAACCCCGCGCTACCACTGACTGGGTCTGTCGACGACCGG 960  
Db 1267 CGGTGCGCGGCGCTGGGCTGAACCCCGCGCTACCACTGACTGGGTCTGTCGACGACCGG 1326  
QY 961 TTGGTTGGGCAACTCTCTCCGAGCCCGGGCTTTCCGCCAAGCGGACCTTAGAGATTGGGGTG 1020  
Db 1327 TTGGTTGGGCAACTCTCTCCGAGCCCGGGCTTTCCGCCAAGCGGACCTTAGAGATTGGGGTG 1386  
QY 1021 ACCGCGCTGGGAGCGCTGGGCGCTACACTGATGCGACCAACCTCTGGTGAAGCC 1080  
Db 1387 ACCGCGCTGGGAGCGCTGGGCGCTACACTGATGCGACCAACCTCTGGTGAAGCC 1446  
QY 1081 ACAGCCCGAGTGAATGCTGGGCTGCTGGCTTGCCTGATGGGAAACAGAGTCCGAG 1140  
Db 1447 ACAGCCCGAGTGAATGCTGGGCTGCTGGCTTGCCTGATGGGAAACAGAGTCCGAG 1506  
QY 1141 GGGGGCGCGCGCGCGCTGGAGCATGCGCGCTGCGCTGCACTGTCGCGAGGGTGAG 1200  
Db 1507 GGGGGCGCGCGCGCGCTGGAGCATGCGCGCTGCGCTGCACTGTCGCGAGGGTGAG 1566  
QY 1201 GGGAGCGTGAAGTCTGAGCCAGCGCTGACAGGTGAACCGGCACTTCAAGGCTG 1260  
Db 1567 GGGAGCGTGAAGTCTGAGCCAGCGCTGACAGGTGAACCGGCACTTCAAGGCTG 1626  
QY 1261 GTGAGCTGGGGTCCCGGGCGCGGCGAGCCAGTGTGATGTTCCAAATCCAGTACAC 1320  
Db 1627 GTGAGCTGGGGTCCCGGGCGCGGCGAGCCAGTGTGATGTTCCAAATCCAGTACAC 1686  
QY 1321 AGCAGCGAAGATGAGACCTCATCTACCGGATTGTCCAGGCTTCAGGCAACACTTCTG 1380  
Db 1687 AGCAGCGAAGATGAGACCTCATCTACCGGATTGTCCAGGCTTCAGGCAACACTTCTG 1746  
QY 1381 CTGAAGCACTGCTCCCGCGGCGTGAATGACTGCTGCTGCTGCTGTCACCGGCG 1440  
Db 1747 CTGAAGCACTGCTCCCGCGGCGTGAATGACTGCTGCTGCTGCTGTCACCGGCG 1806  
QY 1441 GCTGGGCGCTTGAACCTTCAAGCGGCGACAGAGTGTGGGCTGTGCCCTTTCTCCAGGCTG 1500  
Db 1807 GCTGGGCGCTTGAACCTTCAAGCGGCGACAGAGTGTGGGCTGTGCCCTTTCTCCAGGCTG 1866  
QY 1501 CCGGCGCTGCGCGCTGTCAGAGCGCGCTGACAGGCGGCGGAGGCGGAGCGCTGACGCTG 1560  
Db 1867 CCGGCGCTGCGCGCTGTCAGAGCGCGCTGACAGGCGGCGGAGGCGGAGCGCTGACGCTG 1926  
QY 1561 GCCGTGGGCGGCTGTGCTGTGGCTGACCTTACTGTCTTCACTGTGGCGCTTGTGCTGCTGCG 1620

Db 1927 GCCGTGGGCGGCTGTGCTGTGGCTGCTGCTTACTGTCTTCACTGTGCGCTTGTGCTGCG 1986  
QY 1621 GGCCTGGGCGGCGGGAATGCGCGCTCCCTTCACTGAGCTTCACTGCTTCCAGAC 1680  
Db 1987 GGCCTGGGCGGCGGGAATGCGCGCTCCCTTCACTGAGCTTCACTGCTTCCAGAC 2046  
QY 1681 AATGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740  
Db 2047 AATGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2106  
QY 1741 CAGCGAGCTGCTCTTGAACCTTGAAGTGCAGGCTGCTACGTTATGCGAGGCGCTG 1800  
Db 2107 CAGCGAGCTGCTCTTGAACCTTGAAGTGCAGGCTGCTACGTTATGCGAGGCGCTG 2166  
QY 1801 GAGGAGCTTGGGCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1860  
Db 2167 GAGGAGCTTGGGCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2226  
QY 1861 CCGGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1908  
Db 2227 CCGGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2274

## RESULT 3

ADE07450  
ID ADE07450 standard; DNA; 2364 BP.

AC ADE07450;

DT 29-JAN-2004 (first entry)

DE Novel coding sequence (useful for identifying genetic disorders) #516.

KM novel gene; novel protein; tissue marker; molecular weight marker;

KW chromosome marker; genetic disorder; gene; ds.

OS Unidentified.

PN MO2003054152-A2.

PD 03-JUL-2003.

PF 10-DEC-2002; 2002MO-US039555.

PR 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372381P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J, Ma Y, Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z, Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

CC The invention comprises the amino acid and coding sequences of novel proteins. The DNA and protein sequences of the invention are useful as: markers for tissues in which the corresponding protein is preferentially expressed; as molecular weight markers on gels; as chromosome markers or tags; to identify chromosomes or to map related gene positions; and to



XX		
DE	Human NOV44a coding sequence.	
XX		
KW	Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS	
KM	antiinflammatory; cardiact; haemostatic; neuroprotective; anorectic;	
KM	nootropic; immunosuppressive; osteopathic; antiParkinsonian; cancer;	
KM	antifertility; cerebroprotective; gene therapy; NOVA; NOV; fertility;	
KM	metabolic disorder; diabetes; obesity; infectious disease; anorexia;	
KM	neurodegenerative disease; Alzheimer's disease; Parkinson's disease;	
KM	immune disorder; haematopoietic disorder; cardiovascular disorder;	
KM	bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;	
KW	metabolic syndrome X; waiting disorder; cell differentiation;	
KM	Single nucleotide polymorphism; SNP; cell proliferation; haematopoiesis;	
KW	wound healing; angiogenesis; gene; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	variation	replace(604,G)
FT		/tag= a
FT	variation	/standard_name= "Single nucleotide polymorphism"
FT		replace(664,G)
FT		/tag= b
FT	variation	/standard_name= "Single nucleotide polymorphism"
FT		replace(1123,A)
FT		/tag= c
FT	variation	/standard_name= "Single nucleotide polymorphism"
FT		replace(1489,C)
FT		/tag= d
XX		/standard_name= "Single nucleotide polymorphism"
XX		
PN	WO200272771-A2.	
XX		
XX	19-SEP-2002.	
PD		
PX	08-MAR-2002; 2002WC-US007288.	
PF		
XX		
XX		
PR	08-MAR-2001; 2001US-0274101P.	
PR	08-MAR-2001; 2001US-0274194P.	
PR	08-MAR-2001; 2001US-0274281P.	
PR	08-MAR-2001; 2001US-0274322P.	
PR	09-MAR-2001; 2001US-0274849P.	
PR	12-MAR-2001; 2001US-0275235P.	
PR	13-MAR-2001; 2001US-0275578P.	
PR	13-MAR-2001; 2001US-0275579P.	
PR	13-MAR-2001; 2001US-0275601P.	
PR	14-MAR-2001; 2001US-0276000P.	
PR	16-MAR-2001; 2001US-0276176P.	
PR	19-MAR-2001; 2001US-0276994P.	
PR	20-MAR-2001; 2001US-0277239P.	
PR	20-MAR-2001; 2001US-0277321P.	
PR	20-MAR-2001; 2001US-0277357P.	
PR	20-MAR-2001; 2001US-0277381P.	
PR	21-MAR-2001; 2001US-0277791P.	
PR	22-MAR-2001; 2001US-0277833P.	
PR	23-MAR-2001; 2001US-0278152P.	
PR	26-MAR-2001; 2001US-0278894P.	
PR	27-MAR-2001; 2001US-0278999P.	
PR	27-MAR-2001; 2001US-0279036P.	
PR	28-MAR-2001; 2001US-0279344P.	
PR	30-MAR-2001; 2001US-0279995P.	
PR	30-MAR-2001; 2001US-0280233P.	
PR	02-APR-2001; 2001US-0280802P.	
PR	02-APR-2001; 2001US-0280822P.	
PR	02-APR-2001; 2001US-0280900P.	
PR	04-APR-2001; 2001US-0281194P.	
PR	13-APR-2001; 2001US-0283675P.	
PR	30-APR-2001; 2001US-0287424P.	
PR	02-MAY-2001; 2001US-0288066P.	
PR	03-MAY-2001; 2001US-0288142P.	
PR	03-MAY-2001; 2001US-0288528P.	
PR	15-MAY-2001; 2001US-0291150P.	
PR	16-MAY-2001; 2001US-0291099P.	

PR 16-MAY-2001; 2001US-0291240P.  
 PR 30-MAY-2001; 2001US-0294485P.  
 PR 31-MAY-2001; 2001US-0294889P.  
 PR 31-MAY-2001; 2001US-0294899P.  
 PR 18-JUN-2001; 2001US-0290277P.  
 PR 19-JUN-2001; 2001US-0299303P.  
 PR 19-JUN-2001; 2001US-0299310P.  
 PR 10-JUL-2001; 2001US-0304354P.  
 PR 31-JUL-2001; 2001US-0309198P.  
 PR 16-AUG-2001; 2001US-0312903P.  
 PR 10-SEP-2001; 2001US-0318452P.  
 PR 12-SEP-2001; 2001US-0318770P.  
 PR 27-SEP-2001; 2001US-0325430P.  
 PR 27-SEP-2001; 2001US-0325681P.  
 PR 18-OCT-2001; 2001US-0330380P.  
 PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332722P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338092P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 PR 08-MAR-2002; 2002US-00093463.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;  
 PI Boldo FL, Li L, Zernusen BD, Tichner VT, Gangoli EA, Vernet CM;  
 PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;  
 PI Voss EZ, Malayakar UM, Anderson DW, Patturajan M, Miller CE;  
 PI Tadjipier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;  
 PI Zhong M;  
 PI  
 DR WPI, 2002-732824/79.  
 DR P-PSDB; ABP70142.  
 XX  
 PT New NOXV polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOXV-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
 PT disorders, and asthma.  
 PT  
 PS  
 PS Claim 16, Page 261-262; 619pp; English.  
 PS  
 PS The present invention relates to new isolated proteins (NOXV) and their  
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is  
 CC any number from 1 to 48. The NOXV proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOXV-associated disorder.  
 CC The NOXV coding sequences and proteins are useful for treating,  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorders, haematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOXV  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOXV substances for use in  
 CC therapeutic or diagnostic methods  
 CC  
 XX Sequence 3201 BP; 534 A; 1174 C; 986 G; 507 T; 0 U; 0 Other;  
 SQ  
 Query Match 99.4%; Score 1897; DB 6; Length 3201;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1908; Conservative 0; Mismatches 0; Indels 1; Gaps 1



OS	Homo sapiens.
XX	
PN	WO2003025148-A2.
PD	
XX	27-MAR-2003.
XX	
Pf	19-SEP-2002; 2002MO-US029964.
XX	
PR	19-SEP-2001; 2001US-0323739P.
XX	13-SEP-2002; 2002US-00323739.
PA	(HYSEQ-) HYSEQ INC.
PI	Tang YF, Asundi V, Goodrich RM, Ren F, Zhang J, Zhao QA, Wang J;
PT	Ghoch M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PT	Halley-Vicente D;
XX	
DR	WPI; 2003-354603/33.
XX	P-PSDB; ADI21104.
XX	
PS	New polynucleotides and secreted proteins, useful for treating myeloid or
XX	lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
XX	tissue growth or regeneration, in wound healing, and in tissue repair and
XX	replacement.
XX	
PS	Claim 1; SEQ ID NO 79; 156bp; English.
XX	
CC	The invention relates to an isolated polynucleotide encoding a
CC	polypeptide with biological activity. The polynucleotides and
CC	polypeptides are useful in diagnostics, forensics, gene mapping,
CC	identification of mutations responsible for genetic disorders and other
CC	traits, to assess biodiversity, as nutritional sources or supplements.
CC	The polynucleotides may also be used as molecular weight markers,
CC	chromosome markers or map related gene positions, or as an antigen to
CC	raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC	useful for raising antibodies, as markers for tissues in which the
CC	corresponding polypeptide is expressed, for re-engineering damaged or
CC	diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC	bone cartilage, tendon, ligament and/or nerve tissue growth or
CC	regeneration, in wound healing, in tissue repair and replacement, in
CC	healing of burns, incisions and ulcers, and in treating cancer. The
CC	present sequence represents cDNA encoding a novel human protein.
XX	
SQ	Sequence 2885 BP; 443 A; 989 C; 930 G; 523 T; 0 U; 0 Other;
	Query Match            93.8%;   Score 1789.6; DB 10; Length 2885;
	Best Local Similarity   97.1%;   Pred. No. 6e-306;
	Matches 1853; Conservative   0; Mismatches   4; Indels   51; Gaps   1;
OY	1   ATGGCGCCGCGCTCCTGTGTGTGTGTGTGTGCAGCACTTGTGCCACCGAGGACTGCTG
Db	737   ATGGCGCCGCGCTCCTGTGTGTGTGTGTGCAGCACTTGTGCCACCGAGGACTGCTG
OY	61   TGCGTGTCGCAAGAAGCTTCCGAGTGCCTGACCACTTGTGCCACCGAGGACTGCTG
Db	797   TGCGTGTCGCAAGAAGCTTCCGAGTGCCTGACCACTTGTGCCACCGAGGACTGCTG
OY	121   TTGTGTCCGCCCAACGTGTGAACCGCGGCAAGTGAAGCTTGGCTGTGACAATTATC
Db	857   TTGTGTCCGCCCAACGTGTGAACCGCGGCAAGTGAAGCTTGGCTGTGACAATTATC
OY	181   CAGGCGCTTGGAGGCCCTGTGACTTCCGCAACATGAAGGAACTGTGTGAACCTGACT
Db	917   CAGGCGCTTGGAGGCCCTGTGACTTCCGCAACATGAAGGAACTGTGTGAACCTGACT
OY	241   CGCAATGCATCACCCGCAATTTGGAGCCGCTTTGGAGAACCTTGAGAGCTGTGCTC
Db	977   CGCAATGCATCACCCGCAATTTGGAGCCGCTTTGGAGAACCTTGAGAGCTGTGCTC
OY	301   CTCACCTTGACGGGCAAGAGCTGTGTGAAGCTGTGGACACCGGAGAGCTCCGGAGCCGCTC
Db	1037   CTCACCTTGACGGGCAAGAGCTGTGTGAAGCTGTGGACACCGGAGAGCTCCGGAGCCGCTC

QY	361	AATTCGACGACCTTCATCTCTCAGCGGACACCAAGTGGGGCCGACATCGCGCCGGAGCCTTC	420
Db	1097	AATCTGACGACACCTTCATCTCTCAGCGGACACCAAGCTGGGGCCGACATCGCGCGGAGCCTTC	1156
QY	421	GACGACCTTCCTAGAGAGGCTGGAGGACCTGGACCTGTGCTTACCAACAACCTTCGCGCAGAGTG	480
Db	1157	GACGACCTTCCTAGAGAGGCTGGAGGACCTGGACCTGTGCTTACCAACAACCTTCGCGCAGAGTG	1216
QY	481	CCCTGGGCGGCGATCGGCGCCATGCTGCTCCCTGACACCCCTCAACCTTGACCATTAACCTT	540
Db	1217	CCCTGGGCGGCGATCGGCGCCATGCTGCTCCCTGACACCCCTCAACCTTGACCATTAACCTT	1276
QY	541	ATTGACGACCTGCCCCGAGGGGCGCTTGCGCCAGCTCGAGTCAAGCTTCGCGCTGGACCTC	600
Db	1277	ATTGACGACCTGCCCCGAGGGGCGCTTGCGCCAGCTCGAGTCAAGCTTCGCGCTGGACCTC	1336
QY	601	ACCTCCAAACCGCCTGAGCCACGCTGGCTCCGAGCCCGCTTTCTCTCGTGGGCGGTGATGCA	660
Db	1337	ACCTCCAAACCGCCTGAGCCACGCTGGCTCCGAGCCCGCTTTCTCTCGTGGGCGGTGATGCA	1386
QY	661	GAGGCGCTCTCCGCGCCCCCTGAGTGTGCTAGCTTTTAGCGGGAAACCCCGTCGACCTGCAACTGT	720
Db	1387	-----AACCTCGACCTGCAACTGT-----	1405
QY	721	GAGCTGCTGTGACTGCGGCGGCTTGCGCGCCGAGCAGACCTTGAAACGTGCGCTCCCG	780
Db	1406	GAGCTGCTGTGACTGCGGCGGCTTGCGCGCCGAGCAGACCTTGAAACGTGCGCTCCCG	1465
QY	781	CCCGGCGCTGCGCGCGCGCTACTTCTGGGCAAGTGCCTCGAGGGCGAGTTCTCTGTGACCG	840
Db	1466	CCCGGCGCTGCGCGCGCGCTACTTCTGGGCAAGTGCCTCGAGGGCGAGTTCTCTGTGACCG	1525
QY	841	CCCTCATTTGCGCCGACAGCGACGCGCTCTGGGTGCTGGAAGCGCAGCGGACGACGCTG	900
Db	1526	CCCTCATTTGCGCCGACAGCGACGCGCTCTGGGTGCTGGAAGCGCAGCGGACGACGCTG	1585
QY	901	CGGTGCGGGGCGCTGGGTGACCCCGGCGCTACATGCACTGGGTGCGTCTTGACGACCG	960
Db	1586	CGGTGCGGGGCGCTGGGTGACCCCGGCGCTACATGCACTGGGTGCGTCTTGACGACCG	1645
QY	961	TTGTTGTGGCAACTCCTCCCGAGCCCGGAGCTTTCGCCAAACGGAGCTTAGAGATTGGGGTG	1020
Db	1646	TTGTTGTGGCAACTCCTCCCGAGCCCGGAGCTTTCGCCAAACGGAGCTTAGAGATTGGGGTG	1705
QY	1021	ACCGGCGCTGGGGACGCTGGGGGCTACACTGTCATGCGCACCAACCTGTGCTGTAGAGCC	1080
Db	1706	ACCGGCGCTGGGGACGCTGGGGGCTACACTGTCATGCGCACCAACCTGTGCTGTAGAGCC	1765
QY	1081	ACACGCCGAGTATGAACCTGCGGGGTGCTGGCCCTTGCCCATGTTGGGAAACAGAGTGGCGAG	1140
Db	1766	ACACGCCGAGTATGAACCTGCGGGGTGCTGGCCCTTGCCCATGTTGGGAAACAGAGTGGCGAG	1825
QY	1141	GGGGGCGCGCCCGGGGCGCTCGGACATGCGCGCTCCGCTCGCACTGCTGCGAGGGTGAG	1200
Db	1826	GGGGGCGCGCCCGGGGCGCTCGGACATGCGCGCGCTCCGCTCGCACTGCTGCGAGGGTGAG	1885
QY	1201	GGGACGCTGGAAGTTGAGGCCAGCGGTGCAAGTGAACGAGGTGATCGGCACCTCAGGGGCTG	1260
Db	1886	GGGACGCTGGAAGTTGAGGCCAGCGGTGCAAGTGAACGAGGTGATCGGCACCTCAGGGGCTG	1945
QY	1261	GTCAGCTGGGGTCCCGGGCGGCGCAGCGACCAAGTGTGATGTTCCTCAATCCAGTACAAC	1320
Db	1346	GTCAGCTGGGGTCCCGGGCGGCGCAGCGACCAAGTGTGATGTTCCTCAATCCAGTACAAC	2005
QY	1321	AGCAGCGAAGATGAGACCTTCATCTACCGGATTTGTCCAGCGCTTCAGCCACCACTTCTCTG	1380
Db	2006	AGCAGCGAAGATGAGACCTTCATCTACCGGATTTGTCCAGCGCTTCAGCCACCACTTCTCTG	2065
QY	1381	CTGAAGACACTCGTCCCGGGCGCTGACTATAGACCTTCGCTGCGGGCTTTGTACCGGGCC	1440
Db	2066	CTGAAGACACTCGTCCCGGGCGCTGACTATAGACCTTCGCTGCGGGCTTTGTACCGGGCC	2125
QY	1441	GCTGGGGCGCTTCTGACCTCAAGGCGCACCAAGGCTGCTGGGCTGTGCCCACTTCTCCACGCTG	1500









QY 901 CGGTGCGGGGCTGGGTGACCCCGGCTTACATGACTGGTGGTCTTGAACGACCGG 960  
DB 901 CGGTGCGGGGCTGGGTGACCCCGGCTTACATGACTGGTGGTCTTGAACGACCGG 960  
QY 961 TTGTTGGGCACTCTCCCGAGGCGGGGCTTTCGCCAAGGAGACTTAGAGATTGGGGTG 1020  
DB 961 TTGTTGGGCACTCTCCCGAGGCGGGGCTTTCGCCAAGGAGACTTAGAGATTGGGGTG 1020  
QY 1021 ACCGCGCTGGGGGACGCTGGGGGCTACACCTGATGCGACCAACCTTGTGTGAGGCC 1080  
DB 1021 ACCGCGCTGGGGGACGCTGGGGGCTACACCTGATGCGACCAACCTTGTGTGAGGCC 1080  
QY 1081 ACGAGCCGAGTGAATGAGTGGGGGCTGGTGGCTTGGCCCAATGGTGGAAAGAGAGTGGCGAG 1140  
DB 1081 ACGAGCCGAGTGAATGAGTGGGGGCTGGTGGCTTGGCCCAATGGTGGAAAGAGAGTGGCGAG 1140  
QY 1141 GGGGGGCGGGGCGGGGCTTGGGACATGCGCGCTCCGCTGGCACTGCTCCGAGGGTGAG 1200  
DB 1141 GGGGGGCGGGGCGGGGCTTGGGACATGCGCGCTCCGCTGGCACTGCTCCGAGGGTGAG 1200  
QY 1201 GGGAGCGCTGAGTCTGAGCCAGCGGTGAGAGTGAAGAGAGTGAAGAGAGTGAAGAG 1260  
DB 1201 GGGAGCGCTGAGTCTGAGCCAGCGGTGAGAGTGAAGAGAGTGAAGAGAGTGAAGAG 1260  
QY 1261 GTGAGCTGGGGTCCCGGGGCGGCGAGCGGAGTGTGATGATGATGATGATGATGATGATG 1320  
DB 1261 GTGAGCTGGGGTCCCGGGGCGGCGAGCGGAGTGTGATGATGATGATGATGATGATGATG 1320  
QY 1321 AGCAGGAGAGATGAGACCTCATCTACCGGATTTGCTCCAGCGCTCCAGCAGCATTCTCTG 1380  
DB 1321 AGCAGGAGAGATGAGACCTCATCTACCGGATTTGCTCCAGCGCTCCAGCAGCATTCTCTG 1380  
QY 1381 CTGAGACACTCTCTCTCCCGGCGCTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
DB 1381 CTGAGACACTCTCTCTCCCGGCGCTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
QY 1441 GCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1441 GCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 CCGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
DB 1501 CCGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1561 GCGGTGGGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618  
DB 1561 GCGGTGGGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618  
QY 1619 GGGGCGGGGGGGCGGAAATGGCGGCTTCCGCC 1651  
DB 1621 GGGGCGGGGGGGCGGAAATGGCGGCTTCCGCC 1653

## RESULT 8

ABV99421 standard; DNA; 2451 BP.

XX AC ABV99421;  
XX 27-JAN-2003 (first entry)  
DE Human NOV44b coding sequence.  
XX  
XX Human; anti-HIV; cytostatic; antidiabetic; antiaesthetic; cachexia; AIDS;  
XX antiinflammatory; cardiact; haemostatic; neuroprotective; anorectic;  
XX neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;  
XX antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
XX metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
XX immune disorder; haematopoietic disorder; cardiovascular disorder;  
XX bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
XX metabolic syndrome X; wasting disorder; cell differentiation; gene;

KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.  
XX Homo sapiens.  
OS WO200272771-A2.  
PN 19-SEP-2002.  
PF 08-MAR-2002; 2002WO-US007288.  
XX 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
XX 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
XX 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
XX 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
XX 14-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
XX 16-MAR-2001; 2001US-0276767P.  
PR 16-MAR-2001; 2001US-0276994P.  
XX 19-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
XX 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277381P.  
XX 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
XX 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
XX 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279995P.  
XX 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
XX 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
XX 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
XX 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
XX 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
XX 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
XX 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
XX 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
XX 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
XX 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
XX 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
XX 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
XX 18-OCT-2001; 2001US-033080P.  
PR 31-OCT-2001; 2001US-0335301P.  
XX 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332272P.  
XX 14-NOV-2001; 2001US-0332727P.  
PR 14-NOV-2001; 2001US-0333184P.  
XX 14-NOV-2001; 2001US-033372P.  
PR 21-NOV-2001; 2001US-0332094P.  
XX 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
XX 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
XX 08-MAR-2002; 2002US-00093463.

XX (CURA-) CURAGEN CORP.  
 XX  
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;  
 PI Boldog FI, Li L, Zehnen BD, Tchernev VT, Gangoli EA, Vernet CM;  
 PI Pena EA, Burgess CE, Liu X, Spytek KA, Gorman L, Spederna SK;  
 PI Voss EY, Malynkar UM, Anderson DM, Paturajan M, Miller CE;  
 PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Fochart PF;  
 PI Zhong M;  
 XX  
 DR MPI; 2002-732824/79.  
 XX P-SDB; ABP70143.  
 PT New NOVX polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
 PT disorders, and asthma.  
 XX  
 PS Claim 16; Page 262-263; 61pp; English.  
 XX  
 CC The present invention relates to new isolated proteins (NOVX) and their  
 CC coding sequences (ABV9327-ABV9595 and ABP70049-ABP70149), where X is  
 CC any number from 1 to 48. The NOVX proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The NOVX coding sequences and proteins are useful for treating,  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorders, hematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOVX  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods  
 XX  
 SQ Sequence 2451 BP; 363 A; 903 C; 767 G; 418 T; 0 U; 0 Other;  
 Query Match 83.0%; Score 1584.4; DB 6; Length 2451;  
 Best Local Similarity 99.9%; Pred. No. 7.9e-270;  
 Matches 1585; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 421 GACGACTTCTAGAGAGGCTGAGAGACCTGACCTGTCTTACACAACTCCGGCAGGTG 480  
 DB 891 GACGACTTCTAGAGAGGCTGAGAGACCTGACCTGTCTTACACAACTCCGGCAGGTG 950  
 QY 481 CCTGAGCGCGAGATGCGGCGCAGTGCCTGCGACACCTCAACTGAGCACTAACCCTT 540  
 DB 951 CCTGAGCGCGAGATGCGGCGCAGTGCCTGCGACACCTCAACTGAGCACTAACCCTT 1010  
 QY 541 ATTGACGACCTGCCCCAGGCGCTTGGCCAGCTTCAGCTCTCCGCTGACCTC 600  
 DB 1011 ATTGACGACCTGCCCCAGGCGCTTGGCCAGCTTCAGCTCTCCGCTGACCTC 1070  
 QY 601 ACCCTCCACCGGCTGCGGACAGTGCCTGCGACCCGCTTTCTCTGTGGGCTGATGCA 660  
 DB 1071 ACCCTCCACCGGCTGCGGACAGTGCCTGCGACCCGCTTTCTCTGTGGGCTGATGCA 1130  
 QY 661 GAGGCTCTCCGCGCCCTGCTGAGCTTATGAGGGAAACCCCTGCACTGCAACTGT 720  
 DB 1131 GAGGCTCTCCGCGCCCTGCTGAGCTTATGAGGGAAACCCCTGCACTGCAACTGT 1190  
 QY 721 GAGCTGCTGTGCTGCGGCGCTGCGGCGGAGACGACTGAAACGTGCGCTCCCG 780  
 DB 1191 GAGCTGCTGTGCTGCGGCGCTGCGGCGGAGACGACTGAAACGTGCGCTCCCG 1250  
 QY 781 CCGGCGCTGCGGCGGCGCTGCTGAGGAGGAGGCGGAGATTTCTCTGTGAGCCG 840  
 DB 1251 CCGGCGCTGCGGCGGCGCTGCTGAGGAGGAGGCGGAGATTTCTCTGTGAGCCG 1310  
 QY 841 CCGCTCATTTGCGCGGCGGCGGCTGCTGAGGAGGAGGCGGAGATTTCTCTGTGAGCCG 900  
 DB 1311 CCGCTCATTTGCGCGGCGGCGGCTGCTGAGGAGGAGGCGGAGATTTCTCTGTGAGCCG 1370  
 QY 901 CCGTGTCCGGGCGCTGCGGAGTACCCGCGCTTACATGCACTGAGTGGTCTTGACACCGG 960  
 DB 1371 CCGTGTCCGGGCGCTGCGGAGTACCCGCGCTTACATGCACTGAGTGGTCTTGACACCGG 1430  
 QY 961 TTGGTTGGCACTCCTCCCGAGCCCGGCTTTCCCAACGGGACCTTGAAGATTGGGGTG 1020  
 DB 1431 TTGGTTGGCACTCCTCCCGAGCCCGGCTTTCCCAACGGGACCTTGAAGATTGGGGTG 1490  
 QY 1021 ACCGCGCTGAGGAGAGCTGCGGAGGCTTACCTGATCGGCAACCTGCTGTGAGGCG 1080  
 DB 1491 ACCGCGCTGAGGAGAGCTGCGGAGGCTTACCTGATCGGCAACCTGCTGTGAGGCG 1550  
 QY 1081 ACAAGCCGAGTGAAGTGTGCGGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
 DB 1551 ACAAGCCGAGTGAAGTGTGCGGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1610  
 QY 1141 GGGGCGCGCCCGGCGGCTTGGAGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 DB 1611 GGGGCGCGCCCGGCGGCTTGGAGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1670  
 QY 1201 GGAAGCTGAGTGTGAGGAGAGCGGCTGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 1260  
 DB 1671 GGAAGCTGAGTGTGAGGAGAGCGGCTGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 1730  
 QY 1261 GTGAGCTGAGGCTCCCGGCGGCGGAGCGGAGCCAGTGTGATGTTCCAAATCAGTAAAC 1320  
 DB 1731 GTGAGCTGAGGCTCCCGGCGGCGGAGCGGAGCCAGTGTGATGTTCCAAATCAGTAAAC 1790  
 QY 1321 AGCAGGAGATGAGACCTCATATCTACCGGATTTGCCAGGCTGACGACCACTTCTCTG 1380  
 DB 1791 AGCAGGAGATGAGACCTCATATCTACCGGATTTGCCAGGCTGACGACCACTTCTCTG 1850  
 QY 1381 CTGAAGACCTGCTCCCGGCGGCTGACTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 DB 1851 CTGAAGACCTGCTCCCGGCGGCTGACTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1910  
 QY 1441 GCTGGGCGCTTGAACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500  
 DB 1911 GCTGGGCGCTTGAACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1970

QY 1501 CCGGCTTGGCCCTTGTGCAAGCCCTGCAAGCCCACTGCTGGAGGAGGACCTTGAACGTTG 1560  
 DB 1971 CCGGCTTGGCCCTTGTGCAAGCCCTGCAAGCCCACTGCTGGAGGAGGACCTTGAACGTTG 2030  
 QY 1561 GCCGTGGAGGAGTGTGCTGATGAGCTGC 1586  
 DB 2031 GCCGTGGAGGAGTGTGCTGATGAGCTGC 2056  
 RESULT 9  
 AAF76853  
 ID AAF76853 strand: cDNA; 1872 BP.  
 AC AAF76853;  
 XX 14-MAY-2001 (first entry)  
 DE Human secreted protein CDNA #11.  
 XX  
 XX Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW immunosuppressive; antiinflammatory; anti-HIV; immunostimulant;  
 KW cytostatic; cardiac; vascular; anti-angiogenic; ophthalmological;  
 KW neuroprotectant; nocotropic; anticomvulsant; vaccine; antialzheimers;  
 KW antiparkinsonian; antimicrobial; vulnary; gene therapy;  
 KW immune disorder; hyperproliferative; cardiovascular; angiogenic;  
 KW neurological; infection; ss.  
 XX  
 OS Homo sapiens.  
 PN WO200112776-A2.  
 XX  
 PD 22-FEB-2001.  
 PF 15-AUG-2000; 2000WO-US022350.  
 PR 16-AUG-1999; 99US-0148759P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;  
 DR P-PSDB; AAB70072.  
 PT Nucleic acids encoding 18 human secreted polypeptides, useful for  
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
 PT and diabetic retinopathy.  
 XX  
 PS Claim 1; Page 354-355; 380pp; English.  
 CC The present sequence is one of 18 nucleic acid molecules encoding novel  
 CC human secreted proteins. The nucleic acids and proteins may be used in  
 CC the prevention, diagnosis and treatment of diseases including immune  
 CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and  
 CC human immunodeficiency virus (HIV) infections), hyperproliferative  
 CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases  
 CC (e.g. scimitar syndrome, Chaga's cardiomyopathy and coronary  
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft  
 CC neovascularisation and diabetic retinopathy), neurological disorders  
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
 CC infectious diseases and/or for promoting wound healing, regeneration  
 CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar nucleic acid sequences in samples. The polypeptides may also be  
 CC used as antigens in the production of antibodies and in assays to  
 CC identify modulators of protein expression and activity  
 CC  
 SO Sequence 1872 BP; 290 A; 659 C; 604 G; 319 T; 0 U; 0 Other;  
 Query Match 81.9%; Score 1563.4; DB 4; Length 1872;  
 Best Local Similarity 99.9%; Pred. No. 3.9e-266;  
 Matches 1564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 344 GCGTCCGGGGGCGCCGTCATCTGACAGCACTCATCTCAGCGGCAACGAGTGGGCGCA 403  
 DB 1 GCGTCCGGGGGCGCCGTCATCTGACAGCACTCATCTCAGCGGCAACGAGTGGGCGCA 60  
 QY 404 TCGCGCCGGGAGCCTTGCAGCACTTCTTAAGAGCTGAGAGACCTGACCTGTCTTCA 463  
 DB 61 TCGCGCCGGGAGCCTTGCAGCACTTCTTAAGAGAGCTGAGAGACCTGAGACCTGTCTTCA 120  
 QY 464 ACAACCTCCGAGAGTGGCCCTGGGCGGGCATGGGAGCATGGCTGCGGCAACCTTCA 523  
 DB 121 ACAACCTCCGAGAGTGGCCCTGGGCGGGCATGGGAGCATGGCTGCGGCAACCTTCA 180  
 QY 524 ACCTGGAACATTAACCTTAATGACGACTGCCCGGAGGCGCTTGCAGAGCTGGTGAAC 583  
 DB 181 ACCTGGAACATTAACCTTAATGACGACTGCCCGGAGGCGCTTGCAGAGCTGGTGAAC 240  
 QY 584 TCTCCGCGCTGAGACTCACTTCGAAACGCGCTGGCCACGCTGGCTCCGAGCCGCTTTTCT 643  
 DB 241 TCTCCGCGCTGAGACTCACTTCGAAACGCGCTGGCCACGCTGGCTCCGAGCCGCTTTTCT 300  
 QY 644 CTGCGGGGCGTGAATGACAGAGGCTCTCCCGGCCCTGGTGTGAGCTTTAGCGGAAAC 703  
 DB 301 CTGCGGGGCGTGAATGACAGAGGCTCTCCCGGCCCTGGTGTGAGCTTTAGCGGAAAC 360  
 QY 704 CCTGCACTGCAACTGTGAGCTGTGTGAGCTGGCGGCGCTGGCGGCGGACGACTTG 763  
 DB 361 CCTGCACTGCAACTGTGAGCTGTGTGAGCTGGCGGCGGCGGCGGACGACTTG 420  
 QY 764 AAACGTGGGCTTCCCG 823  
 DB 421 AAACGTGGGCTTCCCG 480  
 QY 824 AGTTCTCTGTGAGCG 883  
 DB 481 AGTTCTCTGTGAGCG 540  
 QY 884 GCCAGCGGGCAACGCTGCGGTGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 943  
 DB 541 GCCAGCGGGCAACGCTGCGGTGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 600  
 QY 944 TCGGTCTTGAAGACCGGTGTTGGTGGCACTCTTCCGAGACCGGCGGCTTCCCAAGGGA 1003  
 DB 601 TCGGTCTTGAAGACCGGTGTTGGTGGCACTCTTCCGAGACCGGCGGCTTCCCAAGGGA 660  
 QY 1004 CCTTGAAGATTGGGGTGAACCGGCGCTGGGAGCGTGGGGGCTTACACTTGCATGCCACCA 1063  
 DB 661 CCTTGAAGATTGGGGTGAACCGGCGCTGGGAGCGTGGGGGCTTACACTTGCATGCCACCA 720  
 QY 1064 ACCCTGCTGTGAGGCGCACAGCCCGAGTGAACCTGCGGCTGTGCTGCTTGGCCATGTTG 1123  
 DB 721 ACCCTGCTGTGAGGCGCACAGCCCGAGTGAACCTGCGGCTGTGCTGCTTGGCCATGTTG 780  
 QY 1124 GGAACGACGATGCCGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1183  
 DB 781 GGAACGACGATGCCGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
 QY 1184 CTGCTCCGAGGGTGAAGGAGGAGCGCTGAGTCTGAGCCAGCCGTGCAAGTGAACGAGGTGA 1243  
 DB 841 CTGCTCCGAGGGTGAAGGAGGAGCGCTGAGTCTGAGCCAGCCGTGCAAGTGAACGAGGTGA 900  
 QY 1244 CCGCACTTGAAGGCTGTGAGAGTGGGAGTCCGGGAGGCGGACCGGAGCTGATGTG 1303  
 DB 901 CCGCACTTGAAGGCTGTGAGAGTGGGAGTCCGGGAGGCGGACCGGAGCTGATGTG 960  
 QY 1304 TCCAAATTCAGTACACAGAGGAGAGTGAAGCCGATCTTACCGGATTTGCCAGGCT 1363  
 DB 961 TCCAAATTCAGTACACAGAGGAGAGTGAAGCCGATCTTACCGGATTTGCCAGGCT 1020  
 QY 1364 CCAAGCAACATCTTCTGTGAAGCACTGTCCCGGCGTGAATGAATCTTGTGCTTGC 1423  
 DB 1021 CCAAGCAACATCTTCTGTGAAGCACTGTCCCGGCGTGAATGAATCTTGTGCTTGC 1080  
 QY 1424 TGGCTTTGTCACCGGCGGCTGGGCGCTTGAATCTTACGGCCACAGGCTGTGGCTGTG 1483



Db 1081 TGCCCTTGTACCGGCGGCTGCGGCTCTGACTTCAACGACACAGCTGCTGGGCTGTG 1140  
QY 1484 CCCATTTCCTACAGCTGCGGCTCGCCCTGTGACACGCGCCCTGACAGGCGCAAGTCTGG 1543  
Db 1141 CCCATTTCCTACAGCTGCGGCTCGCCCTGTGACACGCGCCCTGACAGGCGCAAGTCTGG 1200  
QY 1544 GCGGACCCCTGACCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1603  
Db 1201 GCGGACCCCTGACCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
QY 1604 TGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1663  
Db 1261 TGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
QY 1664 AGCTCCAGTCCCAAGACCAATGAGAGGCGCCGACGCGCCCAACGCGCGCGCGGCA 1723  
Db 1321 AGCTCCAGTCCCAAGACCAATGAGAGGCGCCGACGCGCCCAACGCGCGCGCGGCA 1380  
QY 1724 GCGGCGGCG 1783  
Db 1381 GCGGCGGCG 1440  
QY 1784 GTTATGACGAGCGCTGAGGAGGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1843  
Db 1441 GTTATGACGAGCGCTGAGGAGGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
QY 1844 TGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1903  
Db 1501 TGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
QY 1904 TGCTA 1908  
Db 1561 TGTGA 1565

RESULT 10  
ID AAS69014 standard; cDNA; 1582 BP.  
XX AAS69014;  
AC AAS69014;  
XX 13-FEB-2002 (first entry)  
XX DNA encoding novel human diagnostic protein #4818.  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
OS Homo sapiens.  
XX W0200175067-A2.  
XX 11-OCT-2001.  
XX 30-MAR-2001; 2001MO-US008631.  
XX 31-MAR-2000; 2000US-00540217.  
XX 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX P-PSDB; ABG04827.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.  
PS Claim 1; SEQ ID NO 4818; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1582 BP; 216 A; 603 C; 483 G; 279 T; 0 U; 1 Other;

Query Match 72.3%; Score 1379.4; DB 5; Length 1582;  
Best Local Similarity 92.0%; Fred. No. 9.5e-234;

Matches 1560; Conservative 0; Mismatches 12; Indels 124; Gaps 5;

QY 1 ATGCCCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
Db 1 ATGCCCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
QY 61 TCGCTTGGCGAAGCTGTGCGAGTGTGCGACCTCTGTGCGCAACGAGCGCTGCTG 120  
Db 61 TCGCTTGGCGAAGCTGTGCGAGTGTGCGACCTCTGTGCGCAACGAGCGCTGCTG 120  
QY 121 TTTGTGCGCGCGCAACGCTGAGACCGCGCGCAAGTGAAGCTGCGCTGCTGCACTTATC 180  
Db 121 TTTGTGCGCGCGCAACGCTGAGACCGCGCGCAAGTGAAGCTGCGCTGCTGCACTTATC 180  
QY 181 CAGGCGCTGAGGAGCGCGCTGAGCTTCCGCAATGACGCGGACCTGACACTGTCT 240  
Db 181 CAGGCGCTGAGGAGCGCGCTGAGCTTCCGCAATGACGCGGACCTGACACTGTCT 240  
QY 241 CGCAATGCAATCAACCGCATTTGGGCGCGCGCTTTGGGAGCTCGAGAGCTGCTTCC 300  
Db 241 CGCAATGCAATCAACCGCATTTGGGCGCGCGCTTTGGGAGCTCGAGAGCTGCTTCC 300  
QY 301 CTCACCTTGAACGCGCA-CAGGCTGTGAGCTGAGGACG--GGAGCTTCGGGAGCC 357  
Db 301 CTCACCTTGAACGCGCA-CAGGCTGTGAGCTGAGGACG--GGAGCTTCGGGAGCC 357  
QY 358 GTCAATCTGAGACCTCATCTCAGGCGCAACAGCTGAGGCGCATTCGGGAGGCC 417  
Db 358 GTCAATCTGAGACCTCATCTCAGGCGCAACAGCTGAGGCGCATTCGGGAGGCC 417  
QY 418 TTGAGAGACTTCTTAAGAGGCTGAGAGACTGAGACTGCTGTCAACAACTTCGGAG 477  
Db 418 TTGAGAGACTTCTTAAGAGGCTGAGAGACTGAGACTGCTGTCAACAACTTCGGAG 477  
QY 478 GTGCGCTGAGGCGGAGTGTGCGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537  
Db 478 GTGCGCTGAGGCGGAGTGTGCGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537  
QY 538 CTTATTTAGAGCACTGCGCGCGCGCTT-CGCGCACTGTGAGCTGCTGCGCGCTGGA 596  
Db 538 CTTATTTAGAGCACTGCGCGCGCGCTT-CGCGCACTGTGAGCTGCTGCGCGCTGGA 596  
QY 597 CTTATTTAGAGCACTGCGCGCGCGCTT-CGCGCACTGTGAGCTGCTGCGCGCTGGA 600  
Db 597 CTTATTTAGAGCACTGCGCGCGCGCTT-CGCGCACTGTGAGCTGCTGCGCGCTGGA 600  
QY 601 CCTCACCTCAACCGCTGAGCGAGCTGAGCTGCGGACCGGCTTTCTCGTGGGCTGA 660  
Db 601 CCTCACCTCAACCGCTGAGCGAGCTGAGCTGCGGACCGGCTTTCTCGTGGGCTGA 660



Db 121 TTGTNCNCNAAYGTNGAYMGMNAGCTNGARYTNMGTNTGNCNAYAYTTATH 180  
 QY 181 CAGGCCCCGGGGCCCCCTGACTTCCCAACATGACGGGACTGTGTGACCTGACACTGTCT 240  
 Db 181 CAGGCTNTGNCNCNCGNAYTTMGMNAYATGACNGAYTNTGNTGNAAYTTNACNTNMWEN 240  
 QY 241 CGCAATGCATCAACCGGCAATGGGGGCGGCTTTGGGGAACCTCGAGAGCTTCGCTTC 300  
 Db 241 MGNBAVGCAATHACNNGNAHGGNGCMNGCNYTTGGNGAYTNGARFMSNTYTMNGMSEN 300  
 QY 301 CTCACCTTGACGGCAACGAGCTGTGTGAGCTGGGCAACGGGAGCTCCGGGGCCCCGTC 360  
 Db 301 YTNCAAYTNGAYGNAAYMGNVTNTGNTNGARYTNNGNACNGMWSNTYTMNGNCGN 360  
 QY 361 AATCTGCAGACCTCTATCTCTCAGCGCAACCAAGCTGGGCGCATCGCGCGGAGCTTC 420  
 Db 361 AAYYTNACACAYATYNAHYTMWSNGNAAAYCARYTNTGNGMNAHGCNCGNCGNTTY 420  
 QY 421 GACGACTTCTAGAGAGCTGGAGACCTGGACCTGTCTCTACAACAACCTCCGCAAGTG 480  
 Db 421 GAYGAYTTTGTGAMSNYTNNGARAYTNGAYTNTWSNTYAAAYATYTMNGCARGTN 480  
 QY 481 CCTGGGGCGGACATGGGGCCATGCTGCTGCTGCAACCTGACCTGGAACCTTACCTT 540  
 Db 481 CCGTGGGCGNATHTGNGCNAATGCGNGNTNCAACNTYNAAYTTNGAYTATAAYTN 540  
 QY 541 ATTGACGACTGCCCTCCGAGGGGCTTCGCCCAAGCTCGGTGAGCTTCCGCGCTGAGCTC 600  
 Db 541 ATHGAYGNTYTNCCNCGNGGNCNTTYGNCARYTNTGNCARYTWSMNGYTN 600  
 QY 601 ACCTCCACACGCGCTGGGACAGCTGGCTCGGACCCGCTTTCTCTCGTGGGCGTGATCA 660  
 Db 601 ACMSNAAYMGNVTNCAACNTYTNCGNCAAYCCNTYNTYWSMNGMNGMNGAYG 660  
 QY 661 GAGGCTCTCCGCCCCCTGCTGCTGAGCTTTAGCGGAAACCCCTGCACTGCAACTGT 720  
 Db 661 GARGCMNSNCCNGCCNCTYTNMTSNTYTNMNGNAAAYCCNTYTCATGTAAAYTY 720  
 QY 721 GAGCTGCTGTGCTGGCGGCGCTGGCGGCGGACGACGACTGGAAGAGTGCGCTCCCG 780  
 Db 721 GARTYTNNTGTGTNMGMNYTNMGNMNCCNAYGAYTNGARACNTGTGCMWSN 780  
 QY 781 CCGGCGCTGGCGGCGCTGCTTCTGTGGCAATGCCCCGAGGCGGCTTCTCTGTGACCG 840  
 Db 781 CCGNGNTTNGCNGMNTATYTTGTGCGNTNCCNARGNGARITYWSNTGYARCN 840  
 QY 841 CCCCCTCATTTGCCGCAACGACAGGCTCTGAGTGCTGGAAGGCGGCGGACGCTG 900  
 Db 841 CCNTYTNATHGCMNGNCAACNCAARMGNTYTGCTNTYTNARGNCAARMGNCNA 900  
 QY 901 CGGTGGCGGCGCTGGGTGACCCCGGCTTACCATGCACTGGGTGCTCTGACGACG 960  
 Db 901 MGNITGNGCNYTNNGNAGYCCNGCNCNAATGCAATGGGTGNGCCNAGAYAMGN 960  
 QY 961 TTGGTTGGCAACTCTCTCCGAGCCCGGCTTCCCAACGGGACCTTGAAGATTGGGCTG 1020  
 Db 961 YTMGTNGNAAAYMSNMSMNGCMNGCNYTTCNAAYGAMCNYTNGARHTHGAGTN 1020  
 QY 1021 ACCGGCGCTGGGAGCGCTGGGAGCTACACTGCACTGCAACCAACCTGTGCTGAGG 1080  
 Db 1021 ACNGGNGCNGNGAYGNGNGNTYACNTGYATHGCAACNAAYCCNGCNGNAGAR 1080  
 QY 1081 ACAGCCCGAGTAACTGCGGAGTGCTGGGCTTGGCCCATGTTGGGAACAGCACTGCGAG 1140  
 Db 1081 ACNGCMNGNTNGARYTNMNGNTYTNCCNCAAYGNGNAAAYMSMNGNCGAR 1140  
 QY 1141 GGGGGCGGCGGCGCTTGGACATCGCGGCTCCGCTGCACTGCTGCCAGGCTGAG 1200  
 Db 1141 GGGGNGMNGCNGNCGNMSNGAYATHGNGCNGMNGMNGMNGMNGCNGNCGNARGNGAR 1200  
 QY 1201 GGGACGCTGAGTCTGAGCGGCGCTGAGGCTGACGAGGCTGACCGCTGACGAGGCTG 1260  
 Db 1201 GGNACVYTNGARFWSNGARCCNGCNGTNGARGTNAACGAGTNAACNCAACMWSNGN 1260

QY 1261 GTGAGCTGGGCTCCCGGGGCGCCAGCCCAAGTGTGATGTTCCAAATCCAGTACAC 1320  
 Db 1261 GTMSNTGGGNCNCGMNGNCNCGNCCNGTNTGATGTTTCARATHCARAYAAV 1320  
 QY 1321 ACAGCGAAGATGAGACCTCTACCGGATTTGCCAGCTCCAGCCACACTTCCG 1380  
 Db 1321 MSNWSNGARAYGABACNTYNAHTHAYMNAHGTNCCNGCMWSMNCAYCATYTYTN 1380  
 QY 1381 CTGAAGCACTGTCTCCCGGCGCTGACTATGACCTGTGCTGCTGCTGTGTCAACGG 1440  
 Db 1381 YTNARCAAYTNTGTCNCGNGCNGAYTAYGAYTNTGYTNTYTNMGNTMNSNCGN 1440  
 QY 1441 GCTGGGCTCTGACTCTACCGCCACGAGCTGCTGGCTGTGCCATTTCTTCCAGCTG 1500  
 Db 1441 GCGNCCMSNGAYTNAACNCAACNNGNTYTNNGNTGVCNCAATYWSNACNTN 1500  
 QY 1501 CCGGCTCGCCCTGTGTCACGCGCTGAGGCGCAAGCTGTGGGGGAGACCTGACGCTG 1560  
 Db 1501 CCGCMNSNCCNTYTNATGCAAGCTYTNARGNCNAYGNTYTNNGNACNTNACNTN 1560  
 QY 1561 GCGGTGGGGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618  
 Db 1561 GCGTNGGNGGNTYTNGTGNGCNGNTYTNACNGNTYTNCAATYTGCMCTGTGCTTY 1620  
 QY 1619 GGGGCGGGGGCGGGAATAGCGCGCTCCG 1650  
 Db 1621 GNGCNGGNGGNCNAGARATGCGNGCMWSNCC 1652

## RESULT 12

AA166986

ID AA166986 standard; cDNA; 1887 BP.

XX AA166986;

DT 30-JAN-2002 (first entry)

DE Human LRR family member, 33395 coding sequence.

XX LRR; leucine rich repeat; 33395; cytosolic; anti-HIV; antidiabetic;

XX antiarthritic; neuroprotective; dermatological; immunosuppressive; ss;

XX antiinflammatory; antislathmatic; antitumor; antianaemic; hepatotropic.

OS Homo sapiens.

FH Key

FT CDS

FT Location/Qualifiers

FT 1..1887

FT /\*tag= a

FT /product= "33395"

PN W0200172827-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US009470.

PR 24-MAR-2000; 2000US-0191863P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Glucksmann MA;

DR WPI; 2001-626254/72.

XX P-PDB; AAG65805.

New polypeptide for preventing or treating disorders associated with  
 cellular adhesion, proliferation or differentiation, comprises  
 polypeptide 33395, a member of the leucine rich repeat protein family.  
 Claim 1, Fig 1; 133p; English.  
 The invention provides an isolated nucleic acid encoding a polypeptide of













Db 1098 CATTGCGGCATATGAGCTGGCGAGGCCACAGCTGCTGTGAGACTGACTGTGGGTCCCC 1157  
QY 1104 -----GCTGACCTTGCCCATAGTGGGAAACAGCAGTGCAGAGGAGGAGCGGCC 1151  
Db 1158 ACCACCTCCTCAGACTAGCCAACAGCACAGCTGTGACCCCGCGGGAAGGGATCTGA 1217  
QY 1152 CGGAGCCTCGAGCATGCGCGCTCCGCTCGCACTGCTGCCAGAGGTGAGGGAGCGCTGA 1211  
Db 1218 TGCTTCACCCCACTCCGCTGCTCTCTCTTCAGAGTGGCGACATGGGCCCC 1277  
QY 1212 GTCTGAGCCAGCCTGTCAGAGTGAACGAGCCTCAGAGGCTGTGAGCTGGAG 1271  
Db 1278 TACCGACGTTGAGGTCTCAGTATGAGACAGGAGCACAGCTGCTTTGTCAAGTGGC 1337  
QY 1272 TCCCGGCGGCAAGCCGAGCCAGTGTGAATGTTCCAAATCCAGTACAAAGCAGCGAGA 1331  
Db 1338 GGATCAGCGGCTATCCCGGATCCGATGATCCAGATCCAGTACAGTACAGTCCGCTGA 1397  
QY 1332 TGAGACCTCATCTACCGGATTTGTCCAGCTTCAGCCACCACTTCTGTGAGACACT 1391  
Db 1398 TGACATCTCTGTTACAGAGTATCCGCGGAGAGCCGCTGTTCTGTGACGACCT 1457  
QY 1392 CGTCCCGGCGCTGACTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1451  
Db 1458 GAGCTCAGGCGGACCTAGATCTGTGCTGTGCTGTGCTGTATGAGAC--AGCGCAC 1514  
QY 1452 TGACTTCAGGCGCACAGGCTGTGAGCTGTGCCATTTCTCAGCTGCGGCTCGCC 1511  
Db 1515 GGGGCTCAGGCGCACAGGCTGTGAGCTGTGCGCTTCTCCAGCCGAACTGCGCTGCG 1574  
QY 1512 CTTGAGCAGGCGCTGACAGGCGCCAGTGTGAGGCGGACCTGACCGTGGCGGTGGAGG 1571  
Db 1575 GGCATGCGGCGCGCGCTCCCTTCTGTGAGGCGACGATGATCATCGCGTGGGCG 1634  
QY 1572 TGTGCTGAGTGTGCTTACTGATCTTCACTGTGAGCTTGTGAGTTC----- 1618  
Db 1635 CGTATGTAAGCTCGGTAAGTGTCTTCTTCTGTGCTTAATGCGCTAACAGGTGCA 1694  
QY 1619 --GGGCGCGGAGGCGCGAAATGAGCGGCTTCCCTCAGCTCAGCCAGTCCAGTCCA 1676  
Db 1695 CGGCGGCGAGCCCCCGGCAAGGCCAAGATTCGCGCGCTGTAGCAGCGTTTGTCTCCA 1754  
QY 1677 GACCAATGAGAGCGCCAGCCCAACCCAGAGCCCAACCGGC 1718  
Db 1755 GACCAACGCGCGCTTGGGCGCCACGCGCCGCGCGCC 1796

Search completed: November 20, 2004, 08:43:10  
Job time : 1338 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 12:18:50 ; Search time 1361 Seconds

(without alignment)  
7575.395 Million cell updates/sec

Title: US-10-071-879-9

Perfect score: 1908

Sequence: 1 atgagcccccgcgcctctctctc.....tggaaagagatgtggtgtga 1908

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1897	99.4	3201	US-10-093-463-187	Sequence 187, App
2	1787.6	93.7	2563	US-10-093-463-191	Sequence 191, App
3	1629	85.4	1653	US-09-897-214-7	Sequence 7, Appli
4	1584.4	83.0	2451	US-10-093-463-189	Sequence 189, App
5	1563.4	81.9	1872	US-09-768-826-21	Sequence 21, Appli
6	1101.8	57.7	1653	US-09-897-214-9	Sequence 9, Appli
7	624.8	32.7	1887	US-09-815-626-3	Sequence 3, Appli
8	624.8	32.7	1887	US-09-732-436-5	Sequence 5, Appli
9	624.8	32.7	1887	US-10-145-586-62	Sequence 62, Appli
10	624.8	32.7	1887	US-10-221-097-19	Sequence 19, Appli
11	624.8	32.7	2558	US-09-815-626-1	Sequence 1, Appli
12	624.8	32.7	2558	US-10-145-586-60	Sequence 60, Appli

13	624.8	32.7	2963	US-10-388-838-41	Sequence 41, Appli
14	624.8	32.7	3053	US-10-388-838-39	Sequence 39, Appli
15	624.8	32.7	3035	US-10-388-838-43	Sequence 43, Appli
16	624.8	32.7	3245	US-10-302-172-425	Sequence 425, App
17	621	32.5	3144	US-09-732-436-23	Sequence 23, Appli
18	619.4	32.5	2818	US-10-119-926-25	Sequence 25, Appli
19	619.4	32.5	2818	US-10-037-270-92	Sequence 92, Appli
20	619.4	32.5	2818	US-10-117-722-92	Sequence 92, Appli
21	619.4	32.5	2818	US-10-291-172-73	Sequence 73, Appli
22	619.4	32.5	2818	US-10-221-278-73	Sequence 73, Appli
23	617	32.3	2316	US-10-221-097-8	Sequence 8, Appli
24	617	32.3	2869	US-10-479-284-56	Sequence 56, Appli
25	617	32.3	3386	US-10-004-378A-5	Sequence 5, Appli
26	617	32.3	3568	US-10-004-378A-9	Sequence 9, Appli
27	612	32.1	2855	US-10-245-752-99	Sequence 99, Appli
28	612	32.1	2855	US-10-245-859-99	Sequence 99, Appli
29	612	32.1	2855	US-10-245-103-99	Sequence 99, Appli
30	612	32.1	2855	US-10-245-107-99	Sequence 99, Appli
31	612	32.1	2855	US-10-245-143-99	Sequence 99, Appli
32	612	32.1	2855	US-10-245-171-99	Sequence 99, Appli
33	612	32.1	2855	US-10-245-851-99	Sequence 99, Appli
34	612	32.1	2855	US-10-245-883-99	Sequence 99, Appli
35	612	32.1	2855	US-10-237-535-99	Sequence 99, Appli
36	612	32.1	2855	US-10-238-183-99	Sequence 99, Appli
37	612	32.1	2855	US-10-238-283-99	Sequence 99, Appli
38	612	32.1	2855	US-10-238-370-99	Sequence 99, Appli
39	612	32.1	2855	US-10-245-055-99	Sequence 99, Appli
40	612	32.1	2855	US-10-245-147-99	Sequence 99, Appli
41	612	32.1	2855	US-10-245-730-99	Sequence 99, Appli
42	612	32.1	2855	US-10-245-739-99	Sequence 99, Appli
43	612	32.1	2855	US-10-246-210-99	Sequence 99, Appli
44	612	32.1	2855	US-10-239-196-99	Sequence 99, Appli
45	612	32.1	2855	US-10-243-024-99	Sequence 99, Appli

#### ALIGNMENTS

RESULT 1  
US-10-093-463-187  
; Sequence 187, Application US/10093463  
; Publication No. US20030208039A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Shenoy, Sureesh  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Pochart, Pascal  
; APPLICANT: Zhong, Wei  
; APPLICANT: Raetelli, Luca  
; APPLICANT: Mezes, Peter  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Zernusen, Bryan  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Gangolli, Esna  
; APPLICANT: Verne, Corine  
; APPLICANT: Pena, Carol  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Gorman, Linda  
; APPLICANT: Spederna, Steven  
; APPLICANT: Voss, Edward  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Anderson, David  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Miller, Charles  
; APPLICANT: Taupier, Raymond J. Jr.







```

QY 241 CGCATGTCATACCCCGGATTTGGGGCCCGCGCTTTTGGGAGACTTGAGAGCCTGGCTTCC 300
DB 370 CGCATGTCATACCCCGGATTTGGGGCCCGCGCTTTTGGGAGACTTGAGAGCCTGGCTTCC 429
QY 301 CTTCACCTTTGA CGGCAAGGCTGTGTGAGCTGTGGCAACCGGAGCCTTC CGGAGCCCGCTC 360
DB 430 CTTCACCTTTGA CGGCAAGGCTGTGTGAGCTGTGGCAACCGGAGCCTTC CGGAGCCCGCTC 489
QY 361 AATCTGACGACCTTATCTCTAGCGGCAACCAAGCTGGGGCCGATGCGCGCGGAGACCTTC 420
DB 490 AATCTGACGACCTTATCTCTAGCGGCAACCAAGCTGGGGCCGATGCGCGCGGAGACCTTC 549
QY 421 GACGACTTCTTAGAGAGCTTGAGAGACCTGAGACTGTCTCTCAACAACCTTC CGGAGAGTG 480
DB 550 GACGACTTCTTAGAGAGCTTGAGAGACCTGAGACTGTCTCTCAACAACCTTC CGGAGAGTG 609
QY 481 CCTTGGGCGGCGATCGGCGGCAATGCTGCTGCAACCTTC CAACCTTGACCACTTAACTTT 540
DB 610 CCTTGGGCGGCGATCGGCGGCAATGCTGCTGCAACCTTC CAACCTTGACCACTTAACTTT 669
QY 541 ATTGACGACCTGCCCCAGGCGCTTGGCGCACTTGGGCACTGCTCCCGCGCTGAGACTTC 600
DB 670 ATTGACGACCTGCCCCAGGCGCTTGGCGCACTTGGGCACTGCTCCCGCGCTGAGACTTC 729
QY 601 ACTTCAACCGGCTGGGCAAGCTGTGCTCGGACCCGCTTTTCTCTGTGTGGGCGTATGCA 660
DB 730 ACTTCAACCGGCTGGGCAAGCTGTGCTCGGACCCGCTTTTCTCTGTGTGGGCGTATGCA 789
QY 661 GAGGCTCTCCCGCGCCCTGTGTGCTGAGCTTTTAAAGCGGAAACCTTCGTGACCTGTGT 720
DB 790 GAGGCTCTCCCGCGCCCTGTGTGCTGAGCTTTTAAAGCGGAAACCTTCGTGACCTGTGT 849
QY 721 GAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 850 GAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
QY 781 CCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 910 CCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 969
QY 841 CCCCTATTGCGCGGCAACAGCGAGCGCTTGTGGGTGTGTGAAAGCCAGCGGCGCAAGCTGT 900
DB 970 CCCCTATTGCGCGGCAACAGCGAGCGCTTGTGGGTGTGTGAAAGCCAGCGGCGCAAGCTGT 1029
QY 901 CGGTGCGGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 1030 CGGTGCGGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1089
QY 961 TTGGTTGGCAACTCTCTCCGAGCCCGGCTTTTCCCAACCGGACCTTATGAGATTGGGGTG 1020
DB 1090 TTGGTTGGCAACTCTCTCCGAGCCCGGCTTTTCCCAACCGGACCTTATGAGATTGGGGTG 1149
QY 1021 ACCGCGCTGTGGGAGCGTGTGGGCTTACACTGTGATGCGCACCAACCTTGTGTGTGAGGCC 1080
DB 1150 ACCGCGCTGTGGGAGCGTGTGGGCTTACACTGTGATGCGCACCAACCTTGTGTGTGAGGCC 1209
QY 1081 ACAGCCCGGATTAAGACTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB 1210 ACAGCCCGGATTAAGACTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269
QY 1141 GGGGGCGCGC----- 1149
DB 1270 GGGGGCGCGC----- 1329
QY 1150 -----CCGCGGCGCTTGGACATGCGCGCTTCCGCTGTGCACTGTGCGGAGGTTAGGGAGCG 1206
DB 1330 CCGGCGCGGCGCTTGGACATGCGCGCTTCCGCTGTGCACTGTGCGGAGGTTAGGGAGCG 1389
QY 1207 CTGAGAGCTTAGAGCGAGCGCTGTGAGGTGACGAGGTTGACCGGCACTTACGAGGCTGTGTAGC 1266
DB 1390 CTGAGAGCTTAGAGCGAGCGCTGTGAGGTGACGAGGTTGACCGGCACTTACGAGGCTGTGTAGC 1449
QY 1267 TGGGGTCCCGGGCGGCGAGCCGAGCCAGTGTGATGTTCCAAATTCAGTACACAGCAGCG 1326

```

```

DB 1450 TGGGGTCCCGGGCGGCGAGCCAGCCAGCCAGTGTGATGTTCCAAATTCAGTACACAGCAGC 1509
QY 1327 GAAAGTAGACCTCATCTACCGGATTGTCCCAAGCTTCACAGCCACCACTTCTGTGTAAG 1386
DB 1510 GAAAGTAGACCTCATCTACCGGATTGTCCCAAGCTTCACAGCCACCACTTCTGTGTAAG 1569
QY 1387 CACTGTGTCCCGGCGCTGACTATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1446
DB 1570 CACTGTGTCCCGGCGCTGACTATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1629
QY 1447 CCTTGTGACTTACGCGGCAACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1506
DB 1630 CCTTGTGACTTACGCGGCAACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1689
QY 1507 TCGGCCGATGTGCAAGCCCTGTGAGGCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1566
DB 1680 TCGGCCGATGTGCAAGCCCTGTGAGGCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1749
QY 1567 GGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1626
DB 1750 GGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1809
QY 1627 GGGGGCGGAAATGCGCGCTTCCCGCTCAAGCTACAGCAAGTCCAGTCCAGACCAATGGA 1686
DB 1810 GGGGGCGGAAATGCGCGCTTCCCGCTCAAGCTACAGCAAGTCCAGTCCAGACCAATGGA 1869
QY 1687 GGGCCGAGCCGCAACCCAGGCGCCAGCCGCGGAGGCGCCCGCGCGCGAGCGCC 1746
DB 1870 GGGCCGAGCCGCAACCCAGGCGCCAGCCGCGGAGGCGCCCGCGCGCGAGCGCC 1929
QY 1747 AGCTGTCTCTGAGACTGTGAGATGTCGGGCTGTAGGTTTAAAGCGGCGCTGTGGAGGA 1806
DB 1930 AGCTGTCTCTGAGACTGTGAGATGTCGGGCTGTAGGTTTAAAGCGGCGCTGTGGAGGA 1989
QY 1807 GCTTGGGCGCGGAGGAGCACTGTGTGATGAGGAGGCGCTGTGTGTGTGTGTGTGTGTGTGT 1866
DB 1990 GCTTGGGCGCGGAGGAGCACTGTGTGATGAGGAGGCGCTGTGTGTGTGTGTGTGTGTGT 2049
QY 1867 GTAGAGAG 1874
DB 2050 GTAGAGAG 2057

```

```

RESULT 3
US-09-897-214-7
; Sequence 7, Application US/09897214
; Patent No. US20020076779A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell Scott R.
; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,
; TITLE OR INVENTION: Zlrr8, and Zlrr9
; FILE REFERENCE: 01-27
; CURRENT APPLICATION NUMBER: US/09/897,214
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1653)
US-09-897-214-7

```

```

Query Match 85.4%; Score 1629; DB 9; Length 1653;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

[illegible]

OY	1081	A C A G C C G A G T A G A A C T G C G G G T C T G G C C T T T G G C C C A T G S T G G G A A C A G A G T C C G A G	1140
Db	1081	A C A G C C C A G T A G A A C T G C G G G T C T G G C C T T T G G C C C A T G S T G G G A A C A G A G T C C G A G	1140
OY	1141	G G G G G C C G C C C C G G G A C C C T G G A C A T G C G C C T C C G C T G S C A C T G C T G C C G A G G G T G A G	1200
Db	1141	G G G G G C C G C C C C G G G A C C C T G G A C A T G C G C C T C C G C T G S C A C T G C T G C C G A G G G T G A G	1200
OY	1201	G G G A C G C T G A G T C T G A G C C A G C C G T G C A G G T G A O G A G G T G A C C G C C A C C T T C A G G G C T G	1260
Db	1201	G G G A C G C T G A G T C T G A G C C A G C C G T G C A G G T G A O G A G G T G A C C G C C A C C T T C A G G G C T G	1260
OY	1261	G T G A G C T G G G G T C C C C G G G C G G C C A G C C G A C C C A T G T G A T G T T C C A A T C A G T A C A A C	1320
Db	1261	G T G A G C T G G G G T C C C C G G G C G G C C A G C C G A C C C A T G T G A T G T T C C A A A T C A G T A C A A C	1320
OY	1321	A G C A G C G A A G T A G A G A C C C C A T C A T T A C C G A T T G T C C A G C C T C A C A C A C A T T C C T G	1380
Db	1321	A G C A G C G A A G T A G A G A C C C C A T C A T T A C C G A T T G T C C A G C C T C A C A C A C A T T C C T G	1380
OY	1381	C T G A A G C A C T C T G T C C C C G G C G T A C T A T G A C C T T G C C T G C T G G C C T T T G C A C C G A C	1440
Db	1381	C T G A A G A C A C T C T G T C C C C G G C G T A C T A T G A C C T T G C C T G C C T G C T G G C C T T T G A C C G G C C	1440
OY	1441	G C T G G G C C C T T G A A C C T T C A C G A C C A C A G G C T G T G G G A C T G T G C C A T T T C C A C A C C T G	1500
Db	1441	G C T G G G C C C T T G A A C C T T C A C G A C C A C A G G C T G T G G G A C T G T G C C A T T T C C A C A C C T G	1500
OY	1501	C G G G C C T C G A C C C C T G T G C A C A G G C C C T A G G C C C A C G T G C G G G C G G G A C C C T G A A C C G T G	1560
Db	1501	C C G G C C T C G A C C C C T G T G C A C A G C C C T B A G G C C C A C G T G C T G G G C G G G A C C C T G A A C C G T G	1560
OY	1561	G C C G T G G G G G T G N G T G C T G T G G C T T G - C T T A C T G N C T T C A C T G T G - C C T G C T G T T C	1618
Db	1561	G C C G T G G G G G T G N G T G C T G T G G C T T A C T G T T T C A C T G T G G C C C T G T G T G T T C	1620
OY	1619	G G G G C C G G G G G C C G A A A T G G C G G C C T C C C C	1651
Db	1621	G G G G C C G G G G G C C G A A A T G G C G G C C T C C C C	1653

RESULT 4  
US-10-093-463-189  
Sequence 189, Application US/10093463  
Publication No. US20030208039A1  
GENERAL INFORMATION:  
APPLICANT: Padigan, Muralidhara  
APPLICANT: Shenoy, Suresh  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Gusev, Vladimir  
APPLICANT: Pochardt, Pascal  
APPLICANT: Zhong, Mei  
APPLICANT: Rastelli, Luca  
APPLICANT: Mezes, Peter  
APPLICANT: Smithson, Glenda  
APPLICANT: Guo, Xiaofa  
APPLICANT: Gerlach, Valerie  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Zehusen, Bryan  
APPLICANT: Tcherney, Veilizar  
APPLICANT: Gangoli, Esna  
APPLICANT: Verne, Corine  
APPLICANT: Pena, Carol  
APPLICANT: Burgess, Catherine  
APPLICANT: Liu, Xiaobong  
APPLICANT: Spytek, Kimberly  
APPLICANT: Gorman, Linda  
APPLICANT: Spaderna, Steven  
APPLICANT: Voss, Edward  
APPLICANT: Malyankar, Ufiel  
APPLICANT: Anderson, David





























Sat Nov 20 18:13:49 2004

us-10-071-879-9.rnpb

Page 18

Job time : 1371 secs

---

[illegible]

```

RESULT 1
US-09-799-451-425      : Sequence 425, Application US/09799451
Patent No. 6783869
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yuning
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ. ID NOS: 948
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 425
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1187)..(3070)
US-09-799-451-425

```

165 GAGTGAACCTTATCATCAGAGCCCTGAGCCCTTATCTCCGCAACATGACGGGACTGCT 224  
1384 GGCAGACAACTTTCATGCGCTCCGTGCGCGCGCACTGGCAACATGACAGGCTCTCT 1443  
225 GGAACCTGACACTGTCTGCAATGCCATGACCCGCAATGGGGCGCGGCTTTGGGAGACT 284  
1444 GCATCTGAGCTTGTGCGGAAACACATCGGCCACGTGCTGCGGCGCTTTGGCGGACT 1503  
285 CGAGAGCTGCGTTCCTTCCACTTGAACGCAACAGCTGTGAGAGCTGGAGGACGGGAG 344  
1504 GGGGGCCCTGCGTGCCTTGCACCTGATGAGCAACGGCTGACCTCACTGAGGAGGACA 1563  
345 CTTCCGGGGCCCCCTCACTGACAGCACTCTCTAGCGGCAACAGCTGGGCGGACT 404  
1564 GCTGCGCGCTGTGCACTTGCACCTCACTCAAGCAACACAGCTGGCAGCGCT 1623  
405 CGCGCGGGAGCTTTCAGACCTTCTAGAGAGCTGAGAGACTGACCTGCTTACAA 464  
1624 GGGGAGCGGGCGCTGATGATTTGTGCGAGACACTGAGAGACTGACCTTCTTACAA 1683  
465 CAACCTCCGAGAGTCCCTGAGGCGGCAATCGGCGCATGCTGCTGACACCTTCA 524  
1684 CAACCTCCGAGAGTCCCTGAGGCGGCGCTGAGGCGGCGCAACATCAACAGCTTGG 1743  
525 CCGTGAACCTTATCTTATGAGCACTGCGCGCGCGGCTTCCGCAAGCTGCTGAGT 584  
1744 CTTGACCAACACTGTGCTGCTTCTGTGCGCGCGGCGCTTCTTCCGCTGACAGCT 1803  
585 CTCCCGCTGAGACTCAGCTCCCAACCGCTGAGCAAGCTGAGCTGCGGACCGCTTTTCTC 644  
1804 GGGCGGCTGAGCACTGATGACTTCAACCGCTGACCAACATCCACCGGACCACTGCTTC 1863  
645 TGTGGGCGC-----GTGATGACAGAGCTCTCCGCGCGCTGCTGCTGAGTCT 692  
1864 CGGCTGCGCTGCTGCGAGGCGCGCGGCGCTGCGCGCGCTGCTGCTGCTGCTGCT 1923  
693 TAGCGGGAACCCGCTGCACTGCACTGAGCTGCTGCTGCTGCGCGGCGCTGCGGCGC 752  
1924 TGGCGGGAACCCGCTGCACTGCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1983  
753 GAGCACTGAGAAAGCTGCGCTCCCGCGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCT 812  
1984 GAGCACTGAGAGCTGCGCTCCCGCGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCT 2043  
813 GCGCGAGGCGAGTCTCTGTGAGCGCGCTCCTCATGCTCCCGCACAGCGAGGCTCTG 872  
2044 GGGCGGAGGAGGATTTGTCTGCGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 2103  
873 GGTGTGGAAGGCGAGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932  
2104 TGTGCGCGAGGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2163  
933 CATGCACTGAGTCTGCTGAGCAAGCGGCTTGTGCACTCTCCCGAGCGCGGCTTT 992  
2164 TGTGCGTGTGCTGCTGCTGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2223  
993 CCGCAACGAGGACTTGAAGTTGGGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052  
2224 CCGCAATGAGGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2283  
1053 CATGCGCACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103  
2284 CATTTGGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2343  
1104 -----GCTGCGCTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151  
2344 AACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2403  
1152 CGGCGCTTGTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211  
2404 TGTCTTCAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2463

1212 GTCTGAGCAGCCCGTGGAGGTGAGCGAGGTGACACCGCACTCAGGCGTGTGAGCTGAGG 1271  
2464 TACGACCTGTGAGCGTGCAGGTGATGAGCAAGGAGCAAGCTGCTTGTGCACTGAGCC 2523  
1272 TCCCGGCGCGCGCAGCGGACCCAGTGTGATGTTCCAAATCCAGTACAAACAGCAGGAAGA 1331  
2524 GAGTCAAGCGGCTTATCCCGGCGATCCGATGTACAGATTCAGTACAAACAGCTCGGCTGA 2583  
1332 TGAGACCTCATCTTACCGGATTTGTCCAGCTCCAGCCACCACTTCCGCTGGAACACTCT 1391  
2584 TGACATCTCTCTTACAGATGATCCGCGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2643  
1392 GCTCCCGCGCGCTGACTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451  
2644 GCGCTCAGCGCGGACTTACATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
1452 TGACCTCAAGCGCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511  
2701 GGGGCTCAGCGCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
1512 CCTGTCACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1571  
2761 GCGATGCGGCGCGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
1572 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1618  
2821 CCGTATCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1676  
1619 --GGGCGCGGCGCGCGGAAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1676  
2881 CGGCG 2940  
1677 GACCAATGAGGCG 1718  
2941 GACCAAGCGCGCTGCG 2982

RESULT 2  
US-09-620-312D-92  
Sequence 92, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyun  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yundong  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and  
FILE REFERENCE: Polypeptides  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt\_Fl\_genes Version 1.0  
SEQ ID NO 92  
LENGTH: 2818  
TYPE: DNA





```

; TYPE: DNA
; ORGANISM: Homo sapiens
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2019)
US-09-063-950-3

```

Query Match	5.1%;	Score 98.2;	DB 3;	Length 2019;
Best Local Similarity	52.0%;	Pred. No. 5e-11;		
Matches 270; Conservative	0;	Mismatches 243;	Indels 6;	Gaps 2;

QY	17	CCGCCCCCTCTGCTGCTGCTGCTGCGCAGATGAGAGGAGCGCGCTGGCCCTCCCTGAGTGC	66
Db	22	CTGCTGCGGTGTCTCTGCTACTGTGGCCCTGGGGGCTGGGGGTGAGAGGTGCTCATCCGGC	81
QY	67	TGCGAAGACTGTGTCCGATGTGCTCAGACCTCTGTGTCCACCGAAGGCTGTCTTTGTG	128
Db	82	TGCGAATGCG---AGCCAGCCACAGACAGTCTTCTGTCACTGGCCCGCAGGGAGCCACGAGT	138
QY	127	CCGCCCCAAGTGTGAGACGGCGGCAACAGTGGAGAGTGGGGCTGGGCTGACCAATTATCCAGGCC	186
Db	139	CCCCGAGAGCTGTGCACCCGACACGGTGGGGGCTGTACGCTCTTTGAGAGCGGATACCATG	198
QY	187	CTGGGGCCCCCTGACTTTCGCAACATGACGGAGCTGTGTGACCTTGACACTGTCTTCGCAAT	246
Db	199	CTCGAGCGCAGGACAGCTTTTGTCCGGGCTGGCGGGCTGACAGCTCTTGACCTGTCAACAAGC	258
QY	247	GCCATCACCGCAGATTGGGGGCGGGCGCTTTTGGGGGACCTCGAGAGCTGTGGCTTCCCTCCAC	306
Db	259	CAGATGCCAGCTGTGGCCACGGGGGTCTTCCAGCCACTGGCCAACTGACAACTGGAC	318
QY	307	CTTTGACGGCAACAGGCTGTGTGGAGCTGTGGCACCGGAGACTTCCGGGGCCCGCTCAATCTG	366
Db	319	CTGACGGCCAAACAGGCTGTGCATGAATAATCAACATGAGACTTTCGGTGGGCTCGGGCGCTTC	378
QY	367	CAGCACTCATCTCTCAAGCGGCAACCAAGCTGGGCGGCGCATTCGGCGGGAGGCTTGAGAGAC	426
Db	379	GAGCGGCTCTTACTGTGGCAAGAACCGGCAATCGGCACATCCAGCTCGTGCTT---GAGC	438
QY	427	TTTCTTAGAGAGCTGTGAGAGACTTGACCTGTCTTCAACAACTTCGGACAGTGCCTTGG	486
Db	436	AGGCTGACCGGCTCTCTGAGGCTCAAGAGCTGACGAGCAACGAGCTGGGGCACTGCCCG	498
QY	487	GCGGAGATCGGGCCCATGTGCTGCTCCTTGACACACCTCAAC	525
Db	496	CTGGGCTTGCCGCTGTGCTGTGAGACTTCAGCCAC	534

```

RESULT 4
US-09-063-950-1
; Sequence 1, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2852
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(2178)
; IS-09-063-950-1

```

Query Match	5.1%;	Score 98.2;	DB 3;	Length 2852;
Best Local Similarity	52.0%;	Pred. No. 5.3e-11;		
Matches 270;	Conservative	0;	Mismatches 243;	Indels 6; Gaps 2

Qy	7	CCGCGCCTCTGTGCTGTGTGTGCGCAGTGAAGGCGCGCTCCGCGCTGCTGTC	66
Db	181	CTGCTGCGCTGCTCTGCTACTGCGCCCTGGGGCTCGGGGTGCAAGGCTGCCATCCGGC	240
Qy	67	TGCAGAACCTTCCGAGTGCCTGACACCTCTGTGCCACCGAGGCTGTGCTTTGTG	126
Db	241	TGCGATGTC---AGCGAGCCACAGACAGTCTTTGTGCATGTGCCCGCCAGAGGACCAACGATG	297
Qy	127	CCGCGCCAAAGTGAACCGGCGCAGTGTGAGCTGCGGGCTGAGTGAACAATTATCAGAGCC	186
Db	298	CCCCGAGAAGTGCACCCGACACGGTGGGGCTGTAACTCTTGAAGACGGATACCATG	357
Qy	187	CTGGGGCCCCCTGACTTCCGAAATGACGGAGCTGTGACCTTGACATGTCTCGCAAT	246
Db	358	CTCGAGCGAGAGAGCTTTGCCCCGCTCCGGGCTCGACGCTCTGGAACCTGTCAAGAAC	417
Qy	247	GCCATCAACCGGATTTGGGGGCCGGGCTTTTGGGAGCCCTGAGAGCCAGCGTTCCTCCAC	306
Db	418	CAGATTGCCAGCTCTGCCACGCGGGTCTTCCAGCCACTGCGMAACTGACMACTTGGAC	477
Qy	307	CTTGAACGGCAACAGCTGTGTGAGCTGTGGCAACCGAGACCTCCGGGCCCCGTCATCTG	366
Db	478	CTGACGGCCAAAGGCTGTGATGAATTCACATATGAGACTTCCGTGGCTGTGCGGCGCTC	537
Qy	367	CACACCTCATCTCAACGCGGCAACAGACTGGGCGGCAATCGGCGGGAGCGCTTCGAGAC	426
Db	538	GAAAGGCTTACTTGGGCAAGAACCGCATCCGCCACATTCAGCTTGGTCCCTT---GCAC	594
Qy	427	TTCTTGAAGAGCTGAGAGACTTGACCTGTCTTACAACAACCTCCGACAGTGTCCCTG	486
Db	595	AGCCTGACCGGCTCTGTGAGCTCAAGCTGCGAGAGCAACGAGTGTGGGGCACTGCCCG	654
Qy	487	GCCGAGATGGGGCAATGCTGTCCCTTGACACCTCCAC	525
Db	655	CTGGGCTGTCCCGGCTGTGCTGTGGAACCTCAAGCCAC	693

```

RESULT 5
US-09-706-594-6
: Sequence 6, Application US/09706594
: Patent No. 6670123
: GENERAL INFORMATION:
: APPLICANT: New York Blood Center, Inc.
: APPLICANT: Belyavsky, Alexander
: APPLICANT: Shmelkov, Sergei
: APPLICANT: Visser, Jan
: TITLE OF INVENTION: METHOD FOR DETECTING HEMATOPOIETIC STEM CELLS
: FILE REFERENCE: 63475/250
: CURRENT APPLICATION NUMBER: US/09/706,594
: CURRENT FILING DATE: 2000-11-03
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 6
: LENGTH: 4250
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-706-594-6

```

Query Match	4.9%;	Score 93.6;	DB 4;	Length 4250;
Best Local Similarity	54.0%;	Pred. No. 4.7e-10;		
Matches 217;	Conservative 0;	Mismatches 179;	Indels 6;	Gaps 1,

Oy	153	GGAGCTCGGGCTGGCTACAACTTCATCAGAGCCCTGGGGGCCCTCACTTCGCAACAT	212
Db	1261	GAAATGTATCTACAGAGAACTTCATCCGCTGTGTGTGTGCGACAGACAGACTTCTCTGAGGC	1320
Oy	213	GACGGACTGCTGGACCTGACACTGTCTCGCAATGCCATACCCGCACTTGGGGCCGCGC	272
Db	1321	CACGGGCTGGACTCTCGCACTGGGGGAATTAACCGCATCTCGATGATCCAGGACCCGCGC	1380
Oy	273	CTTTGGGGACTCGAGAGCTCGTCTGCTTCCTCCACTTGAAGGCAACAGCGCTGTGTGAGCT	332

Db 1381 TTTCGGGGATCTCACCACCTTAGGGCGCTCTAAGTGCACAGAGTGCAGAGCT 1440  
QY 333 GGGACACGGGAGCCTCCGGGCGCCCGCTCATCTGCAGACCTCATCTCAGCGACCA 392  
Db 1441 GAGCCCGGAGTATTATGATGAGCTGCAGAGCTGCAGATCTCTTCTTCATACATCT 1500  
QY 393 GCTGGGCGCATCGCGCGGAGCCTTGCAGACATCTCTAGAGAGCCTGCAGAGCCTGA 452  
Db 1501 CATCCGCGAGATCTGAGTCTGAACTTTGACCCGGTCCCAACCTCC-----AGCTGCT 1554  
QY 453 CCTGCTCTACACACCTCCGAGGTGCTGCGCGCGCATCGGCGCATCTGCGCT 512  
Db 1555 ATCTGTAATACAACTCTCTCAGGCGCATGCTTCAGGCGTCTTCTTGCGCTTACCT 1614  
QY 513 GCACACCTCACTGACCTGACCATCTTATGACGACCTGCC 554  
Db 1615 CCTCAGGCTAAACCTGAGAGTAAACCATCTTCACTCTCTTCC 1656

RESULT 6  
US-08-592-500-1  
; Sequence 1, Application US/08592500  
; Patent No. 6005089  
; GENERAL INFORMATION:  
; APPLICANT: Lanza, Francois  
; APPLICANT: Phillips, David R.  
; APPLICANT: Cazenave, Jean-Pierre  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,500  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/089,455  
; FILING DATE: 09-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 12418-28  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7452 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: 1462..2419  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2422..4101  
; FEATURE:  
; NAME/KEY: misc signal  
; LOCATION: 68..76

OTHER INFORMATION: //function= "Putitive TPA responsive  
OTHER INFORMATION: element"  
OTHER INFORMATION: /label= TRE  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 471..478  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: repeat\_region  
LOCATION: 593..881  
OTHER INFORMATION: //tpc\_type= "other"  
OTHER INFORMATION: /label= Alu  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 897..904  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1142..1149  
OTHER INFORMATION: //function= "Sp1 binding site"  
OTHER INFORMATION: /standard\_name= "Sp1"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 1178..1184  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequences"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1199..1203  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1263..1269  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1285..1289  
OTHER INFORMATION: //function= "GATA-1 binding site"  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1321..1326  
OTHER INFORMATION: //function= "GATA-1 binding site"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 1365..1372  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequences"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: repeat\_region  
LOCATION: 6133..6440  
OTHER INFORMATION: //tpc\_type= "other"  
OTHER INFORMATION: /label= Alu  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 5610..5615  
OTHER INFORMATION: //standard\_name= "Polyadenylation  
OTHER INFORMATION: signal sequence"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 6966..6971  
OTHER INFORMATION: //standard\_name= "Polyadenylation  
OTHER INFORMATION: signal sequence"  
FEATURE:

```
NAME/KEY: misc_signal
LOCATION: 7224..7229
OTHER INFORMATION: /standard_name="Polyadenylation"
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7358..7363
OTHER INFORMATION: /standard_name="Polyadenylation"
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..7452
OTHER INFORMATION: /standard_name="Nucleotide"
OTHER INFORMATION: sequence containing the human GPV gene"
US-08-592-500-1

Query Match
Best Local Similarity 47.8%; Score 91.2; DB 3; Length 7452;
Matches 264; Conservative 0; Mismatches 288; Indels 0; Gaps 0;

QY 273 CTTTGGGACCTTGAGAGCTGGTCCCTCCACCTTGAGCGCAAGCTGTGAGCT 332
Db 3276 CTTGGGAGATGGGGGCTTGAGAGCTGTGCTGAACCGCACCTGCGACCT 3335
QY 333 GGGACACGGGAGCTCCGGGGCCCGCTCAATGACAGCACTGATCTGAGCGGCAACA 392
Db 3336 GCGCGCGCGCGCTTCCGCACTGAGCGCGCTGCGGTACTTAAAGGGGTGACTGTAGGCC 3395
QY 393 GCTGGGCGCGATCGCGCGGAGCCTTCAAGCACTTCTAAGAGGCTTGAGAGACCTTGA 452
Db 3396 GCGGCTGAGCGGCTTCCGAGGCGCGCTTCCAGAGGCTTGAGAGCTCAGGTGCTCGC 3455
QY 453 CCTGTCTCAACAACCTCCGAGAGTGGCGCGGCGCGATGAGGCGCATGCTGCTCCCT 512
Db 3456 CCTGCACTCCACAGGCTGACCGCTCCCGAGCGCTTGCTGCGCGCTCGGCAAGCT 3515
QY 513 GCACACCTCAACCTGAGACCATTAATTGACGACTGCGCCGAGCGCTTCGCCCA 572
Db 3516 GCGCCAGGAGTCTCTGCGCGCAAGGCTGCGCGCGCTGCGCGCTTCTCCGCA 3575
QY 573 GCTGGTCACTCTCCGCGCTGAGCTTCACTCAACCGCGCTGCGCAAGCTGCTCCGA 632
Db 3576 TCTCAGCAGCTGAGAGGCTCCAGCTCGACCAACAGCTGAGAACCTGCTGCGCA 3635
QY 633 CCGGCTTTCTCGGCGGCTGATGAGAGGCTCTCCCGCGCGCTGCTGAGCTT 692
Db 3636 CGTGTGGGCTCTGCGCGGCTGACGAGAGGCTGTTGGGCGACAACCTCTGCGCTG 3695
QY 693 TAGCGGAAACCCCTGCTGACCTGACAGTGTGCTGCTGCGGCTGCGCGCGCC 752
Db 3696 CGACTGTGCTGCGGCGCTTCTGCGGCTGCTGCGGCGAGCACTTAGGCTGCTGCGG 3755
QY 753 GGAAGACCTGAAAGCTGCGCTTCCCGCGCGCGCTGCGCGCGCTTCTTGGGCAAGT 812
Db 3756 GGAAGAGCCCGCGCTGCGGCGAGGCTTGGGCGCGAGCGCGCGCTGCGCTGCGGCGCT 3815
QY 813 GCCCGAGGCGGA 824
Db 3816 GCCCGAGGCTGA 3827

RESULT 7
US-08-195-006-1
Sequence 1, Application US/08195006
Patent No. 6083688
GENERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
APPLICANT: Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Kourie and Crew
```

```
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,006
FILING DATE: 10-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: intron
LOCATION: 1462..2419
NAME/KEY: CDS
LOCATION: 2422..4101
FEATURE:
NAME/KEY: misc_signal
LOCATION: 68..76
OTHER INFORMATION: /function="Putative TPA responsive"
OTHER INFORMATION: element"
OTHER INFORMATION: /label= TRE
FEATURE:
NAME/KEY: misc_signal
LOCATION: 471..478
OTHER INFORMATION: /function="Ets-1 cis-acting"
OTHER INFORMATION: sequence"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: misc_signal
LOCATION: 493..502
OTHER INFORMATION: /function="Ets-1 cis-acting"
OTHER INFORMATION: sequence"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: repeat_region
LOCATION: 593..881
OTHER INFORMATION: /rpt_type="other"
OTHER INFORMATION: /label= A1u
FEATURE:
NAME/KEY: misc_signal
LOCATION: 897..904
OTHER INFORMATION: /function="Ets-1 cis-acting"
OTHER INFORMATION: sequence"
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1142..1149
OTHER INFORMATION: /function="Spi binding site"
OTHER INFORMATION: /standard_name="Spi1"
```

```

FEATURE:
NAME/KEY: misc_signal
LOCATION: 1178..1184
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequences
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1199..1203
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1263..1269
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1285..1289
OTHER INFORMATION: /function= "GATA-1 binding site"
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1321..1326
OTHER INFORMATION: /function= "GATA-1 binding site"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 1365..1372
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequences
FEATURE:
NAME/KEY: repeat_region
LOCATION: 6133..6440
OTHER INFORMATION: /ftc_type= "other"
OTHER INFORMATION: /label= A1u
FEATURE:
NAME/KEY: misc_signal
LOCATION: 5610..5615
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 6966..6971
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7224..7229
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7358..7363
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..7452
OTHER INFORMATION: /standard_name= "Nucleotide"
OTHER INFORMATION: sequence containing the human GPV gene"
US-08-195-006-1

Query Match 4.8%; Score 91.2; DB 3; Length 7452;
Best Local Similarity 47.8%; Pred. No. 1.6e-09;
Matches 264; Conservative 0; Mismatches 288; Indels 0; Gaps 0;
QY 273 CTTTGGGACCTGAGAGCTGCTCCCTCCACCTTGACCGCAACAGGCTGTGAGCT 332
DB 3276 CTTGGGAGATGGGGGCTTGCAGAGCTGTGGCTGAACCGCACCACTGGCAGCCT 3335
QY 333 GGGCACCAGGAGCTTCGGGGCCCGCTCAATCTGACGACCTCATCTCAGCGGCAACCA 392
DB 3336 GCCCGCGCGCGCTTCGCAACTGAGCGCGCTGCGGTACTTATGGGGTGACTCTGAGCCC 3395
QY 393 GCTGGGCGGATGGCGCGGAGCTTTCAGACGACTTCTTGAAGAGCTGAGGAGCACTTGA 452
DB 3396 GCGGCTGAGCGGCTTTCGAGGCGCTTTCAGGCGCTTGGGAGCTCCAGGTGCTGCG 3455

```

```

QY 453 CTTGCTCTACAAACACTCCGCGAGAGTGCCCTGGGCGGCAATCGGCGCCACTGCGCCT 512
DB 3456 CTTGACTTCGAAGGCGCTGACCGGCGCTCCCGAGCGGCTTGTGCGCGGCTGGGAGCT 3515
QY 513 GCACACCTCAACCTGACGACATTAACCTTATGACGACACTGCGGCGGCGCTTGGCCA 572
DB 3516 GCGCAGGTGTCCTCGCGCGCGGCAACAGGCTGGCGCGCTGCGCGCTTCTTCCGCAA 3575
QY 573 GCTGGGTGAGCTTCCCGCTGAGACTCACTCCGCAACCGCGCGGCAAGCTGCGGCTTCCGCA 632
DB 3576 TCTCAGCAGCTGAGAGAGCTGACAGCTGACCAACAGCTGAGAGAGCTGCGGCGCA 3635
QY 633 CCGGCTTTCTCTGCTGGGCGGATGACAGAGGCTCTCCCGCGCGCTTGTGCTGAGCTT 692
DB 3636 CGTGTGGGCTCTGCGCGGCTGACGAGGCTCTGTGGGCGACACTCTGCGGCTG 3695
QY 693 TAGCGGAAACCCCTGCACTGACACTGAGCTGTGCTGCGGCGGCTGCGGCGG 752
DB 3696 CAGCTGGGCTGGGCGGCTTCTGAGGCTGCTGCGGCGGCACTAGGCTGCTGAGGCG 3755
QY 753 GAGCAGCTGGAAGAGTGGCTCCCGCGCGGCGGCTGCGGCGGCTTCTTGGGAGT 812
DB 3756 GGAAGAGCGGCGGCGGCTGAGGCTGAGGCGGCGGCGGCGGCTGCGGCTGAGGCGCT 3815
QY 813 GCCCGAGGCGGA 824
DB 3816 GCCGCGGCGTGA 3827

RESULT 8
PCT-US94-07644A-1
Sequence 1, Application PC/TUS9407644A
GENERAL INFORMATION:
APPLICANT: COR Therapeutics, Inc.
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07644A
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 012418-003000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: inton
LOCATION: 1462..2419
FEATURE:
NAME/KEY: CDS

```

```

1 LOCATION: 6966..6971
2 OTHER INFORMATION: /standard_name="Polyadenylation
3 OTHER INFORMATION: signal sequence"
4 FEATURE:
5 NAME/KEY: misc_signal
6 LOCATION: 7224..7229
7 OTHER INFORMATION: /standard_name="Polyadenylation
8 OTHER INFORMATION: signal sequence"
9 FEATURE:
10 NAME/KEY: misc_signal
11 LOCATION: 7358..7363
12 OTHER INFORMATION: /standard_name="Polyadenylation
13 OTHER INFORMATION: signal sequence"
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: 1..7452
17 OTHER INFORMATION: /standard_name="Nucleotide
18 OTHER INFORMATION: sequence containing the human GPV gene"
19 PCT-US94-07644A-1

```

Query Match	Similarity	4.8%;	Score 91.2;	DB 5	Length 7452;
Best Local Similarity	47.8%;	Pred. NO. 1.6e-09;			
Matches	264;	Conservative	0;	Mismatches 288;	Indels 0; Gaps 0
QY	273	CTTTGGGAGACTCGAGAGCTTGCCTTCCCTACCTTGAAGGCAACAGACTGGTAAAGCT	332		
Db	3276	CTTTCGGGGAGATGGGGGGGCTCTGAGAGACTGTGGCTGAACGGCAACCGACTGGCAACCT	3333		
QY	333	GGGACACGGGAGACTCCGGGGGCCCCCTCAATCTTGACAGACCTCATCTCAAGGGACCA	392		
Db	3336	GCCCGCCGCGCTTCGCGCAACTGAGCGGCTGCGTACTTAAGGGGAGACTTGAAGCC	3395		
QY	393	GCTTGGGCGCATCGCGCCGGGAGCTTTCGACGACTTCTTAGAGAGCTGGAGGACCTGGA	452		
Db	3396	GGGCTGAGCGCGCTTCCGACAGGGCGCTTTCAGAGGCTTGGCGAGCTCCAGTGTCTGC	3455		
QY	453	CTGTCTCTAACACAACTCCCGGAGGTGCCCTGGGCGCGGACATCGCGCCATGCTGCCCT	512		
Db	3456	CTTGCATCTCCAAACGGGCTGACCGGCTTCCCGACGGGCTTGTGCGGGCTCTGGGCAAGCT	3515		
QY	513	GCACACCTTCAACTGTGAACCATTAACCTTATTGACGCACTGGCCCCAGCGCTTTCGCCA	572		
Db	3516	GGGCGAGGTGCTCCCTGGCGCGCAACAGACTGGCGGCTTGCCTTCCCTTCGCGCA	3575		
QY	573	GCTCGGTCAAGCTTCCCGGCTGAGCTCACTCACTCCAAACGGCTGGGCAACGCTGGCTCGGA	632		
Db	3576	TCTTCAGACGCTTGGAGAGGCTTCAGCTTCGACCAACCAAGTGGAGAACCTGTGCTGGCGA	3639		
QY	633	CCCGCTTTTCTCTGTGGGGGTGATGAGAGAGCTCTCCGCCCCCTGTGCTGAGCTT	692		
Db	3636	CGATTTTGGGGGCTCTGCGCGGCTGACGGAGGTCTCTTGGGGGCAACACTCTGGGCTG	3695		
QY	693	TAGCGGAAACCCCTGCACTGCAACTGTGACTGTGTGCTGGCGGACTGGCGGAGC	752		
Db	3696	CGACTGTGGCTTGGGGGCTTCTCTGGGAGTGGCTGGGCAAGCACTTAAGGCTCTGTGGGGG	3755		
QY	753	GGACGACCTGAGAAAGTGGGCTTCCCGCGCGGCTGGCGGCGGCACTTCTTGGGAGT	812		
Db	3756	GGAGAGACCCCCACAGGTGCGGAGGCTTGGGGCGCAAGCGGCTTGCCGCTTGTGGGCTT	3815		
QY	813	GCCCGAGGGCGA	824		
Db	3816	GCCCGGGGGTGA	3827		



```
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yunding
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pl_FL_genes Version 2.0
SEQ ID NO 480
LENGTH: 2733
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (271)..(2730)
US-09-799-451-480
```

```
Query Match
Best Local Similarity 4.6%; Score 87.8; DB 4; Length 2733;
Matches 252; Conservative 0; Mismatches 252; Indels 3; Gaps 1;
```

```
QY 289 AGCTGCGTTCCTCCACCTTACCGGCAACAGGCTGCTGAGACTGGGCAACCGGAGCCTC 348
DB 436 ACCGGACGACGCTGGGCTCAACGAGCAAGCTCAAGCTGCTTACTCTTCGCTC 495
QY 349 CCGGCCCCCGTCAATGTGAGACCTTCAAGGCGGCAACAGCTGGGCGGCACTGG 408
DB 496 AACCGTTTGGAACTCCACCGACCTCAACCTCAACGAAAGATCTCTTACATCGAG 555
QY 409 CCGGAGCCTTGCAGACCTTCTAGAGAGCTGAGAGACTGGAACCTGTCTACAACAAC 468
DB 556 GAGGCTTCTCTGCG---GCCAGTGAAGCTGAGAGTCTCTCAAGCTGGGCTACAACAAG 612
QY 469 CTCGGCAGGCTGCTGGGCGCGGATCGGCGCCATGCTGCTGCTGCAACCTCAACTG 528
DB 613 CTCAGCAACCTGACGAGGAGGATGCTGCGAGGATAGCGGCTGCAAGTTCTTTTTC 672
QY 529 GACCAATACCTTATTGAGCGACGCGCCGAGGCGCTTGCCAGCTGGTCAAGCTTCC 588
DB 673 CAGCAACCTTCAATCGAGGTGAGCGCCACCGCTTCTCGAGTGCAGAGCTCATC 732
QY 589 CCGCTGAGCCTTCACTCAACCGCTGAGCGAGCTGAGCTGAGACCGCTTTTCTCTCGT 648
DB 733 AGAGTGAAGCTGTCTCCAAACCGCTCAAGCGCTGAGAGGTGCACTTTTGCAAGCTC 792
QY 649 GAGCGTGAAGAGAGCGCTTCCCGCGCTGCTGCTGAGCTTGAAGCGGAAACCCCTG 708
DB 793 GCCAGCTGATGCTGTGTGAGCTGCGCGCAACCCCTTCACTGATGAGCGACTTTC 852
QY 709 CACTGCACTGAGTGTGCTGAGCTGCGGCGGCTGGCGGCGGCGGAGCAACCTGAGAAAG 768
DB 853 GCGTCTCTGAGCTGCGCTGAGGTCTTCAACAACGATCAACAAGAACTACGACGCTGAG 912
QY 769 TGGCGCTCCCGCGCGCGCTGAGCGGCG 795
DB 913 TGTGAGTCCCGCGGAGGATTTGCCGAGC 939
```

RESULT 10

```
US-10-140-002-381
Sequence 381, Application US/10140002
Patent No. 6725730
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C59
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 381
LENGTH: 2236
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-381
```

```
Query Match
Best Local Similarity 4.6%; Score 87.6; DB 4; Length 2236;
Matches 222; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
```

```
QY 181 CAGGCGCTGGGGCCCCCTGACTTCCGCAACATGACGAGGACTGTGAGACTGACACTGTCT 240
DB 649 CAGGCGCTGAGAGCTGTGCGCTGGGCACTCCTGCTGCGAGCGAGCATCTTCTGCAAC 708
QY 241 CGCAATGCGATCACCGGATTTGGGCGCGGCTTTGGGAGCTTGAAGAGCTGCTTCC 300
DB 709 GGCACCGCATCTCGCATGTGCGACCTGCACTTCCGCTGCGCAACTCAACATC 768
QY 301 CTCACCTTGAAGGCAACAGGCTGTGAGAGCTGGGCAACCGGAGCTCCGGGGCCCCCTC 360
DB 769 CTGTGCTGCACTCGAATGTGCTGGCCCGAATTTGATGCGGCTTCACTGAGCTTGGCC 828
QY 361 AATGTCAGACCTTATCTTCAAGCGCAACAGCTGAGCGGATCGCGCGGAGCTTTC 420
DB 829 CTCCTGAGAGCTGAGCTTCAAGCATATGACAGCTCCGCTGTGTGAAACCTGCAACA 888
QY 421 GACGACTTCTAGAGAGCTGAGAGACTGAGCTGTCTCTAACAACAACCTCCGGCAGGTG 480
DB 889 TTCCAGCGCTGGGCGGCTTACACAGCTGCACTGAGACCGCTGCGGCTGAGAGAGCTG 948
QY 481 CCTGGGCGGATGAGGCGGCAATGCTGCGCTGCAACCTCACTGAGACCATACCTT 540
DB 949 GGCCTGGGCTGTCTTCCGCGCTGCTGCTGCTGCTGCTGCTTACTGAGAGCAACGCG 1008
QY 541 ATTGAAGCACTGCGCGGAGCGCTTGGCGAGCTCGGTCAAGCTTCCGCTGAGACTC 600
DB 1009 CTGCAAGCACTGCTGATGACACTTCCGCGACCTGGGCAACCTCAACAACCTTCTCC 1068
QY 601 ACCTCCAAACCGCTGAGCAAGCTGAGC 626
DB 1069 CACGGCAACCGCATCTCCAGCTGCC 1094
```

RESULT 11  
US-09-461-325-40

```

; Sequence 40, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; PRIOR FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-461-325-40

```

```

Query Match
Best Local Similarity 49.6%; Score 86; DB 4; Length 1777;
Matches 221; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

```

```

QY 181 CAGGCGCTGGGGCCCCCTGACTTCGCGAATGACGGAGCTGGTGAACCTGACATGTCT 240
DB 174 CAGGCGCTGAGGCTGTGCGCCGCGGAGCATCTGCTGCGACGCCAGCGCATCTTCTGAC 233
QY 241 CGCAATGCCATCACCAGCATTTGGGGCCCGGCGCTTTGGGGGACCTGAGAGCTGGCTTC 300
DB 234 GCGAACCGCATCTCGCATGTGCCAGCTGCCAGCTTCGCTGCGCAACCTTCCACCATC 293
QY 301 CTCACCTTGAGCGGCAACAGGCTGTGAGAGCTGGGCAACCGGAGCCTCCGGGGCCCCGTC 360
DB 294 CTGTGCTGACACTCGAATGTGTGCGCCGCAATGATGCGGCTGCTTCACTGCGCTGGCC 353
QY 361 AATCTGACAGCACTCTATCTTCAAGCGGCAACAGCTGGGCTGCTGCGCGGAGCTTC 420
DB 354 CTCCTGAGAGCTGAGACCTCAAGCAAGTAAATGCAACAGCTCGGTGTGTGACCTTGCACA 413
QY 421 GAGCACTTCTAGAGAGCTGAGAGCACTGACCTGTGCTTAAACAACCTCCGGGAGGTG 480
DB 414 TTCAAGGCGCTGGGCGGCTTACACAGGTGCACTTGAACCGCTGCGGCTTGCAGAGCTG 473
QY 481 CCTTGGGCGGAGCATCGGCGCATGCTGCTGCTGACACCTTCAACCTGAGCACTTAACCTT 540
DB 474 GGGCCGCGGCTGTTCGCGGCGCTGCTGCTGCTGCACTTCACTGAGAGCAACGCG 533
QY 541 ATTGAGCACTGCGCCGAGCGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 534 CTGCAAGGCACTGCTGATATACACTTTCGCGAGCTGGGCAACCTTCAACACCTCTTCTG 593
QY 601 ACCTCCACCGCTGCGCAAGCTGGC 626
DB 594 CAGCGCAACCGCATCTCCAGCGTGC 619

```

```

RESULT 12
; Sequence 40, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

```

```

; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-012-542-40

```

```

Query Match
Best Local Similarity 49.6%; Score 86; DB 4; Length 1777;
Matches 221; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

```

```

QY 181 CAGGCGCTGGGGCCCCCTGACTTCGCGAATGACGGAGCTGGTGAACCTGACATGTCT 240
DB 174 CAGGCGCTGAGGCTGTGCGCCGCGGAGCATCTGCTGCGACGCCAGCGCATCTTCTGAC 233
QY 241 CGCAATGCCATCACCAGCATTTGGGGCCCGGCGCTTTGGGGGACCTGAGAGCTGGCTTC 300
DB 234 GCGAACCGCATCTCGCATGTGCCAGCTGCCAGCTTCGCTGCGCAACCTTCCACCATC 293
QY 301 CTCACCTTGAGCGGCAACAGGCTGTGAGAGCTGGGCAACCGGAGCCTCCGGGGCCCCGTC 360
DB 294 CTGTGCTGACACTCGAATGTGTGCGCCGCAATGATGCGGCTGCTTCACTGCGCTGGCC 353
QY 361 AATCTGACAGCACTCTATCTTCAAGCGGCAACAGCTGGGCTGCTGCGCGGAGCTTC 420
DB 354 CTCCTGAGAGCTGAGACCTCAAGCAAGTAAATGCAACAGCTCGGTGTGTGACCTTGCACA 413
QY 421 GAGCACTTCTAGAGAGCTGAGAGCACTGACCTGTGCTTAAACAACCTCCGGGAGGTG 480
DB 414 TTCAAGGCGCTGGGCGGCTTACACAGGTGCACTTGAACCGCTGCGGCTTGCAGAGCTG 473
QY 481 CCTTGGGCGGAGCATCGGCGCATGCTGCTGCTGACACCTTCAACCTTGAACCTTAACCTT 540
DB 474 GGGCCGCGGCTGTTCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533
QY 541 ATTGAGCACTGCGCCGAGCGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 534 CTGCAAGGCACTGCTGATATACACTTTCGCGAGCTGGGCAACCTTCAACACCTCTTCTG 593
QY 601 ACCTCCACCGCTGCGCAAGCTGGC 626
DB 594 CAGCGCAACCGCATCTCCAGCGTGC 619

```

```

RESULT 13
; Sequence 40, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; FILE REFERENCE: P2029G30A1D2
; CURRENT APPLICATION NUMBER: US/10/115,123

```

```
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-40
```

```
Query Match 4.5%; Score 86; DB 4; Length 1777;
Best Local Similarity 49.6%; Pred. No. 1.4e-08;
Matches 221; Conservative 0; Mismatches 225; Indels 0; Gaps 0;
```

```
QY 181 CAGGCGCTGGGCGCCCTGACTTCGCAATGACGGGACTGTGAGACCTGACAGTGTCT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 174 CAGGCGCTGGGCGCCCTGACTTCGCAATGACGGGACTGTGAGACCTGACAGTGTCT 233
QY 241 CGCAATGCCATACCGCGATTG3GGCGCGGCTTTGGGGACCTGAGAGCCTGGGTTCC 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 234 GGCACCGCATCTCGCATGTGCGAGCTGCGACGCTTCGCTCCGCGCAACTCACAATC 293
QY 301 CTCACCTTGACGGCAACAGGCTGTGAGCTGGGACCTGGAGAGCTCCGGGGCCCCCTC 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 294 CTGTGGCTGCACTCGAATGTGCTGGCGCCGCAATTGATGCGGCTCTTCACTGGCCCTG 353
QY 361 AATCGACGACCTCATCTCTCAGCGGCAACGAGCTGGGCGGATGGCGCGCGGAGCCTTC 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 354 CTCCTGGAGACCTGAGCTTCAGCGATATGCAACAGCTCCGCTGTGTGAGACCTTGAC 413
QY 421 GACGACTTCTTAGAGAGCTTGAGAGACCTGACCTGTCTCTTACAAACCTTCGGAGAGTG 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 414 TTCCACGCGCTGGCGCGCTTACACAGGTGACCTGGAGACCGCTGGAGAGAGTG 473
QY 481 CCTGGCGCGGATGAGCGCCATGCTGCTGCTGACACCTTCAACTGAGACATACTT 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 474 GCGCGCGGCTGTTCGCGCGCTGTGCTGCTGACATCTTACCTGACAGAGACGCG 533
QY 541 ATTGACGACCTGCGCCGAGGCGCTTGCGCGAGCTGCTGAGCTCTCCGCGCTGAGACTC 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 534 CTGACGAGCACTGCTGATACACCTTCGCGAGCTGGGCAACCTCACACACTTCTCTG 593
QY 601 ACCTCCACCGGCTGGCCACGCTGGC 626
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 594 CACGGCAACCGCATCTCCGCGGTGCC 619
```

```
RESULT 14
US-09-520-781-9
; Sequence 9, Application US/09520781
; Patent No. 6689866
; GENERAL INFORMATION:
; APPLICANT: Shimkels, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 No. 6689866el Polynucleotides
; CURRENT APPLICATION NUMBER: US/09/520,781
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: USSN 60/123,667
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 9
; LENGTH: 2341
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215) .. (2173)
US-09-520-781-9
```

```
Query Match 4.2%; Score 79.6; DB 4; Length 2341;
Best Local Similarity 51.1%; Pred. No. 2.8e-07;
Matches 187; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
```

```
QY 58 CCTGCGTGTGCGCAACCTTGCGAGTGTGAGACCTCTGTGGCCACCGAGGCTG 117
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 353 CCTGCGTGTGCTGTGTGAGTAACCACTTACAGAGAGTGTGTGACCGCGGGGCTTC 412
QY 118 CTGTTGTGCGCGCCCAAGTGTGACCGGCGACAGTGTGAGCTGGCTGTGACAACTTC 177
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 413 TCCGAGTTCGCGAGGATATTCCTCGAACAACCGGTACTCACTCATATGAGAAACAAC 472
QY 178 ATCCAGCGCCCTGGGGCCCCCTGACTTCGCAACATGAGGGACTGTGTGACTGACACTG 237
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 473 ATCCAGATGATCCAGGCGGACACCTTCGCACTTCAACAGTGTGAGAGTGTGAGTTG 532
QY 238 TCTGCAATGCAATCACCGGATTTGGGCGCGGCTTTGGGAGACCTGAGAGCCTGCGT 297
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 533 GCGAGGAATCCATCCGCGAGATTGAGTGTGGGCGCTTCAACGCGCTGGCGAGCTCAGC 592
QY 298 TCCCTCCACTTTCAGCGCAACAGGCTGTGTGAGCTGGGACCCGAGAGCTCCGGGGCCCC 357
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 593 ACCCTGAGAGCTGTGCAACAATGCTGACATCCCTAGCGGGGCTTTGAAATTACTG 652
QY 358 GTCAATCTGAGCACTCATCTCTACGCGGCAACGAGTGTGGCGGATGGCGCGGAGGCG 417
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 653 TCCAGCTGCGGAGCTGTGCTTGTGCAACAACCCATGAAAGATCCCTTTCAGCC 712
QY 418 TTGAC 423
| | | | |
DB 713 TTCAAC 718
```

```
RESULT 15
US-09-520-781-11
; Sequence 11, Application US/09520781
; Patent No. 6689866
; GENERAL INFORMATION:
; APPLICANT: Shimkels, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 No. 6689866el Polynucleotides
; CURRENT APPLICATION NUMBER: US/09/520,781
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: USSN 60/123,667
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2607
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215) .. (1984)
US-09-520-781-11
```

```
Query Match 4.2%; Score 79.6; DB 4; Length 2607;
Best Local Similarity 51.1%; Pred. No. 2.8e-07;
Matches 187; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
```

```
QY 58 CCTGCGTGTGCGCAACCTTGCGAGTGTGAGACCTCTGTGGCCACCGAGGCTG 117
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 353 CCTGCGTGTGCTGTGTGAGTAACCACTTACAGAGAGTGTGTGACCGCGGGGCTTC 412
```

```
QY 118 CTGTTGTGCGGCCCAACGTGSAACCGGCGACAGTGGAGCTGGGCTGAGCAACTTC 177
Db 413 TCCGAGGTCCCGCAGGGTATTCCTCGAACACCCGGTACTCAACCTCATGGAGAACAC 472
QY 178 ATCCAGGCCCTGGGGCCCTCTGACTTCGCAACATGACGGGACTGGTGAAGTGAACACTG 237
Db 473 ATCCAGATGATCCAGGCCGACACCTTCGGCACCTCCACACCTGGAGGTCTGCAAGTTG 532
QY 238 TCTCGCAATGCCATCAACCGGCAATTGGGGGCCGCGCTTTGGGAACTTCGAGAGCTGGCT 297
Db 533 GGCAGGAACCTCCATCCGGCAGATTGAGGTGGGGCCCTTCACGCGCTGGCCAGCTCAGC 592
QY 298 TCCCTCCACCTTGACGGCAACAGGCTGATGAGCTGGGCAACGGGAGCCCTCCGGGGCCC 357
Db 593 ACCCTGGAGCTGTTCGACAACTGGCTGACAGTCAATCCCTAGCGGGGCTTTGAATACCTG 652
QY 358 GTCAATCTGAGCACTCATCTCTAGGGGCAACCAAGCTGGGCCGCAATGGCGCGGAGCC 417
Db 653 TCCAGCTGGGGAGCTTGGCTTCGCAACAAACCCCATGMAAGCATCCCTCTTAAGCC 712
QY 418 TTGAC 423
Db 713 TTCAAC 718
```

Search completed: November 20, 2004, 14:51:20  
Job time : 219 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 06:40:05 ; Search time 12916 Seconds  
(without alignments)  
6985.812 Million cell updates/sec

Title: US-10-071-879-9  
Perfect score: 1908  
Sequence: 1 atgccccgcgcctcctcgtc.....tgaagaagagatgtgtgtga 1908

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hhg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1908	100.0	2449	9	BC015581 Homo sapi
2	1665	87.3	1665	6	CO718542 Sequence
3	1629	85.4	1653	6	AX358336
4	1585.4	83.1	1989	9	BC027475
5	1525.4	79.9	2553	10	BC023036
6	1523.8	79.9	2612	10	BC023156
7	1355	71.0	139072	9	AP000485
8	1355	71.0	182769	2	AC013176
9	1335	69.4	180153	2	AC018798
10	1314	68.9	1944	9	BC007718
11	1185	60.5	1571	9	BC071866
12	1108.2	58.1	238317	2	AC119332
13	1101.8	57.7	1653	6	AX358338
14	1083.8	56.8	269255	2	AC106165
15	1082.2	56.7	195342	10	AC128739
16	992	52.0	1288	9	BC014040
17	985	51.6	1286	9	BC000207
18	624.8	32.7	1867	6	AX179258
19	624.8	32.7	1887	6	AX266997

20	624.8	32.7	2067	9	AK172754	AK172754 Homo sapi
21	624.8	32.7	2549	9	BC003578	BC003578 Homo sapi
22	624.8	32.7	2558	6	AX266995	AX266995 Sequence
23	621	32.5	3144	9	AB033072	AB033072 Homo sapi
24	619.4	32.5	2818	6	AR338601	AR338601 Sequence
25	617.4	32.4	3320	9	AB056799	AB056799 Macaca fa
26	617	32.3	2316	6	CQ730862	CQ730862 Sequence
27	617	32.3	3386	6	AX675281	AX675281 Sequence
28	617	32.3	3568	6	AX675285	AX675285 Sequence
29	615.4	32.3	1734	6	CO721191	CO721191 Sequence
30	612	32.1	2855	6	AX574572	AX574572 Sequence
31	612	32.1	2855	9	AT358127	AT358127 Homo sapi
32	591.2	31.0	2450	6	AX675283	AX675283 Sequence
33	583.4	30.6	2911	10	AK122476	AK122476 Mus muscu
34	552.4	29.0	3174	6	AX676904	AX676904 Sequence
35	552.4	29.0	3174	9	AB040917	AB040917 Homo sapi
36	549.2	28.8	2831	10	BC066999	BC066999 Mus muscu
37	518.8	27.2	138538	9	AC011445	AC011445 Homo sapi
38	518.8	27.2	172090	2	AC018477	AC018477 Homo sapi
39	518.6	27.2	92797	9	AF038458	AF038458 Homo sapi
40	511	26.8	100426	9	AL591063	AL591063 Human DNA
41	488.2	25.6	159443	2	AC101792	AC101792 Mus muscu
42	488.2	25.6	221980	10	AC124583	AC124583 Mus muscu
43	487	25.5	247615	2	AC098061	AC098061 Rattus no
44	487	25.5	277055	2	AC127880	AC127880 Rattus no
45	482.4	25.3	179820	2	AC011679	AC011679 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS  
DEFINITION  
BC015581  
Homo sapiens leucine rich repeat and fibronectin type III domain containing 4, mRNA (cDNA clone MGC:23278 IMAGE:4637053), complete cds.

ACCESSION  
BC015581  
VERSION  
BC015581.2  
KEYWORDS  
MGC.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
1 (bases 1 to 2449)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Wax,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullen,B., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richard,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hilyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickinson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalek,A., Smallus,D.E., Scherch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2449)  
Strausberg,R.  
Direct Submission  
Submitted (04-OCT-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.







Db 961 TTGTTGGCAACTCTCCCGAGCCGGGCTTTCCCAACGGGACCTTAGAGATTGGGGTG 1020  
 QY 1021 ACCGGGCTGGGGAGCGTGGGGGTATACCTGCAATGCCACCAACCTTGCTGTGAGGCC 1080  
 Db 1021 ACCGGGCTGGGGAGCGTGGGGGTATACCTGCAATGCCACCAACCTTGCTGTGAGGCC 1080  
 QY 1081 ACAGCCCGATAGAACTGCGGGTGTCTGGCTTGGCCCATGGTGGGAACAGAGTGGCCAG 1140  
 Db 1081 ACAGCCCGATAGAACTGCGGGTGTCTGGCTTGGCCCATGGTGGGAACAGAGTGGCCAG 1140  
 QY 1141 GGGGGGCGGGCGGGGCTTGGACATGCGGCTTCCCTGTGCACTGCTGCGAGGGTGA 1200  
 Db 1141 GGGGGGCGGGCGGGGCTTGGACATGCGGCTTCCCTGTGCACTGCTGCGAGGGTGA 1200  
 QY 1201 GGGAGCTGGAGTGTGAGCCAGCGGTGAGAGGTGACCGGACCTCAAGGGCTG 1260  
 Db 1201 GGGAGCTGGAGTGTGAGCCAGCGGTGAGAGGTGACCGGACCTCAAGGGCTG 1260  
 QY 1261 GTGAGCTGGGGTCCCGGGCGGGCCAGCCGACCAAGTGTGATGTTCCAAATCCAGTACA 1320  
 Db 1261 GTGAGCTGGGGTCCCGGGCGGGCCAGCCGACCAAGTGTGATGTTCCAAATCCAGTACA 1320  
 QY 1321 AGAGAGGAATAGAACCTCTATCTACCGGATTGTCCAGCTTCCAGCCACACTTCTG 1380  
 Db 1321 AGAGAGGAATAGAACCTCTATCTACCGGATTGTCCAGCTTCCAGCCACACTTCTG 1380  
 QY 1381 CTGAAGCACTCTGCTCCCGGCGGTGACTATGACCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 Db 1381 CTGAAGCACTCTGCTCCCGGCGGTGACTATGACCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 QY 1441 GCTGGGCTCTGACCTCAAGCCACAGGCTGCTGGCTGTGCTGCTGCTGCTGCTGCTG 1500  
 Db 1441 GCTGGGCTCTGACCTCAAGCCACAGGCTGCTGGCTGTGCTGCTGCTGCTGCTGCTG 1500  
 QY 1501 CGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
 Db 1501 CGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
 QY 1561 GCGGTGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 Db 1561 GCGGTGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 QY 1621 GGGCGGGGGGGCGGAAATGCGCGCTCCCTCAAGCTCAAGCCAC 1665  
 Db 1621 GGGCGGGGGGGCGGAAATGCGCGCTCCCTCAAGCTCAAGCCAC 1665

RESULT 3  
 AX358336 1653 bp DNA linear PAT 13-FEB-2002  
 LOCUS AX358336  
 DEFINITION Sequence 7 from Patent WO0202604.  
 ACCESSION AX358336  
 VERSION AX358336.1 GI:18674972  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Thayer, E.C., Sheppard, P.O. and Presnell, S.R.  
 TITLE Leucine-rich repeat proteins, zlr7, zlr8 and zlr9  
 JOURNAL Patent: WO 0202604-A 7 10-JAN-2002;  
 ZymoGenetics, Inc. (US)  
 FEATURES  
 source  
 1..1653  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..>1653  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD23323.1"  
 /db\_xref="GI:18674973"

/translation="MAPPLLILILASGAACPVPVCQNLSESLITCAHRLGLFVPP  
 NVNRIVTELALDNFIQALGPDPRMTGTVIDLTLLRNATITRIGARAFGLBLSLH  
 LDNRILVELGTGSLRGVNLHLILSGNQLGRJAPGAPDFLSLDLIDLSYNNLKV  
 PMAGIAMPALHTLNLIDNLIDALPGAFGLDLRLDLTSRLTLADLPFESGR  
 DAPASAPLVLSFGNDPHCNCLMRLRLRADDLETCSPPGLAGRWYAPDEPF  
 SCEPPLIARHTOBLMYLEGRATLRALDPAPTMYHWYPPDRIVNSSRAAFNG  
 TLEIGTGAADAGGTTCIATNPAGEATARVELVLAIPHGNSSABGRPSDLAS  
 ARTABSEGLTSESPAVQVEVATGSLVSWGGRPADPVMMFOIYNSDEDTLYR  
 ITPASSHHFLKHLVPGADYDLCILALSPAGSDLTATRLCGAHSTLPASPLCHA  
 LQAHVIGTLTVAVGVLVALTLHCGPCWFGAGPEMAASP"

ORIGIN  
 Query Match 85.4%; Score 1629; DB 6; Length 1653;  
 Best Local Similarity 99.9%; Pred. No. 1,1e-216;  
 Matches 1651; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 ATGGCCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
 Db 1 ATGGCCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
 QY 61 TGGCTTGCAGAACCTGTCCGAGTGTGCTGAGCACTCTGTGCCACCGAGGCTGTG 120  
 Db 61 TGGCTTGCAGAACCTGTCCGAGTGTGCTGAGCACTCTGTGCCACCGAGGCTGTG 120  
 QY 121 TTGTGCGCGCCCAACGTTGACCGCGGCAAGTGAAGCTGTGGCTGACAACTTCAATC 180  
 Db 121 TTGTGCGCGCCCAACGTTGACCGCGGCAAGTGAAGCTGTGGCTGACAACTTCAATC 180  
 QY 181 CAGGCCCTGGGGCCCTCTGACTTCCGCAACATGACGGGACTGGTGAACCTGACACTGTCT 240  
 Db 181 CAGGCCCTGGGGCCCTCTGACTTCCGCAACATGACGGGACTGGTGAACCTGACACTGTCT 240  
 QY 241 CGCAATGCCATCAACCCGCAATTGGGCGCGGCTTTGGGGAACCTGAGAGCCCTGCTCC 300  
 Db 241 CGCAATGCCATCAACCCGCAATTGGGCGCGGCTTTGGGGAACCTGAGAGCCCTGCTCC 300  
 QY 301 CTCACCTTGAAGGGAACAGGCTGTGAGCTGTGGGACCGGGAGCTCCGGGGCCCCGTC 360  
 Db 301 CTCACCTTGAAGGGAACAGGCTGTGAGCTGTGGGACCGGGAGCTCCGGGGCCCCGTC 360  
 QY 361 AATCTGCAGACACTATCTCTAGCGGCAACAGCTGGGCGGATCGCGCGGAGGCTTTC 420  
 Db 361 AATCTGCAGACACTATCTCTAGCGGCAACAGCTGGGCGGATCGCGCGGAGGCTTTC 420  
 QY 421 GACGACTTCTTGAAGAGCTTGAAGACCTTGAACCTGTCTTAACAACACTCCGGCAGGTG 480  
 Db 421 GACGACTTCTTGAAGAGCTTGAAGACCTTGAACCTGTCTTAACAACACTCCGGCAGGTG 480  
 QY 481 CCTGGGCGGGCATGCGGCGCATGCTGTGCGCAACCTGCAACCTGGAACATACCTT 540  
 Db 481 CCTGGGCGGGCATGCGGCGCATGCTGTGCGCAACCTGCAACCTGGAACATACCTT 540  
 QY 541 ATTGAAGCACTGCCCCCAAGCGCTTGTGCGCACTGTGCTGCTGCTGCTGCTGCTGCTG 600  
 Db 541 ATTGAAGCACTGCCCCCAAGCGCTTGTGCGCACTGTGCTGCTGCTGCTGCTGCTGCTG 600  
 QY 601 ACCTCAACCGCTGCGGCAAGCTGTGCTGCGAACCGGCTTTCTCTGTGGGCGTGAATGA 660  
 Db 601 ACCTCAACCGCTGCGGCAAGCTGTGCTGCGAACCGGCTTTCTCTGTGGGCGTGAATGA 660  
 QY 661 GAGGCTCTCGCGCGCCCTGTGCTGAGCTTGAAGGGAACCCCTGCACTGCAACTGT 720  
 Db 661 GAGGCTCTCGCGCGCCCTGTGCTGAGCTTGAAGGGAACCCCTGCACTGCAACTGT 720  
 QY 721 GAGTGTGTGTGCTGCGGCGGTGCGGCGGCAAGCACTTGAACCTGTGCGCTTCCCG 780  
 Db 721 GAGTGTGTGTGCTGCGGCGGTGCGGCGGCAAGCACTTGAACCTGTGCGCTTCCCG 780  
 QY 781 CCGGCGCTGGCGCGCGCTTACTTCTTGGGCAATGCGCGGCGAGTTCTCTGTGAGCGG 840  
 Db 781 CCGGCGCTGGCGCGCGCTTACTTCTTGGGCAATGCGCGGCGAGTTCTCTGTGAGCGG 840  
 QY 841 CCGCTCATTTGCGCGGCAACGAGCGCTTGGGCTGTGAAGGCAACGCGGCGCAAGCTG 900

```

Db      |||||
841 CCCCTCATTTGCCGCGACACGACGCGCTCTGGGTGTGGAAAGCCAGCGGCGCACGCTG 900
Qy      |||||
901 CCGTCCCGGCGCCCTGGGTGTACCCCGGCGCTACCAATGCACTGGGTGGTCTTGACACCGG 960
Db      |||||
901 CCGTCCCGGCGCCCTGGGTGTACCCCGGCGCTACCAATGCACTGGGTGGTCTTGACACCGG 960
Qy      |||||
961 TTGGTTGGCACTCCCTCCGAGCCCGGGCTTCCCAACGAGGACCTTATGAGATGGGGTG 1020
Db      |||||
961 TTGGTTGGCACTCCCTCCGAGCCCGGGCTTCCCAACGAGGACCTTATGAGATGGGGTG 1020
Qy      |||||
1021 ACCGCGCTGGGAGCGCTGGGGGCTACACCTTCATTCGCCACCAACCTCTGGTGAAGCC 1080
Db      |||||
1021 ACCGCGCTGGGAGCGCTGGGGGCTACACCTTCATTCGCCACCAACCTCTGGTGAAGCC 1080
Qy      |||||
1081 ACAGCCCGAGTGAACCTGGGGGTCTGGCTTCCCATGGTGGGAAACAGCAGTCCGAG 1140
Db      |||||
1081 ACAGCCCGAGTGAACCTGGGGGTCTGGCTTCCCATGGTGGGAAACAGCAGTCCGAG 1140
Qy      |||||
1141 GGGGGCGGCGCCCGGGGCGCTCGGACATCGCGCGCTCGCTCGCACTGGTCCGAGGGGTAG 1200
Db      |||||
1141 GGGGGCGGCGCCCGGGGCGCTCGGACATCGCGCGCTCGCTCGCACTGGTCCGAGGGGTAG 1200
Qy      |||||
1201 GGGAGCGCTGAGTCTGAGCGACCGCTGAGTGAAGGAGTGAACGCGCACTCAGGGCTG 1260
Db      |||||
1201 GGGAGCGCTGAGTCTGAGCGACCGCTGAGTGAAGGAGTGAACGCGCACTCAGGGCTG 1260
Qy      |||||
1261 GTGAGCTGGGGTCCCGGGGCGGCGACCGCAAGTGTGATGTCCAAATCCAGTAAAC 1320
Db      |||||
1261 GTGAGCTGGGGTCCCGGGGCGGCGACCGCAAGTGTGATGTCCAAATCCAGTAAAC 1320
Qy      |||||
1321 AGCAGGAGATGAGACCTCTCACTACCGGATTTGTCCAGGCTCCAGCCACCACTTCTG 1380
Db      |||||
1321 AGCAGGAGATGAGACCTCTCACTACCGGATTTGTCCAGGCTCCAGCCACCACTTCTG 1380
Qy      |||||
1381 CTGAAGCACCTGCTCCCGGCGCTGACTATGACTCTGCTGCTGCTGTGTACCGGCGC 1440
Db      |||||
1381 CTGAAGCACCTGCTCCCGGCGCTGACTATGACTCTGCTGCTGCTGTGTACCGGCGC 1440
Qy      |||||
1441 GCTGGGCGCTCTGACTCTACCGGCGACGACGAGGCTGTGGGCTGTGCCATTTTCCAGCTG 1500
Db      |||||
1441 GCTGGGCGCTCTGACTCTACCGGCGACGACGAGGCTGTGGGCTGTGCCATTTTCCAGCTG 1500
Qy      |||||
1501 CCGGCGCTCGCCCTGTGACCGCGCTGACGAGCCCACTGCTGGGCGGAGCCCTGACCGTG 1560
Db      |||||
1501 CCGGCGCTCGCCCTGTGACCGCGCTGACGAGCCCACTGCTGGGCGGAGCCCTGACCGTG 1560
Qy      |||||
1561 GCCGTGGGGGTGTGTGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618
Db      |||||
1561 GCCGTGGGGGTGTGTGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618
Qy      |||||
1619 GGGGCGGCGGCGGCGGAGAAATGCGGCGCTCCCGC 1651
Db      |||||
1621 GGGGCGGCGGCGGCGGAGAAATGCGGCGCTCCCGC 1653

```

```

RESULT 4
BC027475 1889 bp mRNA linear PRI 04-JUN-2004
LOCUS BC027475
DEFINITION Homo sapiens leucine rich repeat and fibronectin type III domain
ACCESSION BC027475
VERSION containing 4, mRNA (cDNA clone IMAGE:3887234), partial cds.
KEYWORDS BC027475.1 GI:22268152
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS
Strasberg, R.D., Collins, F.S., Wagner, L., Stamen, C.M., Schlier, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F.,

```

```

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshitsugu, S.,
Carninci, P., Prange, C., Raha, S.S., Loguigliano, N.A., Peters, G.J.,
Abramson, R.D., Mullen, S.J., Bosk, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Hellon, E., Kettelman, M., Madan, A., Rodriques, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1889)
Strasberg, R.
Direct Submission
Submitted (04-Apr-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseg, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```

## REMARK

COMMENT

## FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
Series: IRAP Plate: 52 Row: a Column: 22.

gene

CDS

```

1..1889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3887234"
/tissue_type="lung, carcinoma, large cell
undifferentiated."
/clone_1b="NIH_MGC_69"
/lab_host="DH10B"
/note="vector: PCMV-SPORE6"
<1..1889
/gene="LRFN4"
/note="synonym: MGC3103"
/db_xref="LocusID:78999"
<1..1587
/gene="LRFN4"
/codon_start=1
/product="LRFN4 protein"
/protein_id="AAH27475.2"
/db_xref="GI:48257123"
/db_xref="LocusID:78999"
/tranlation="LVECTGSLRGVNLQHLILSGNOLGRIAPGAFDDFLESLIEDD
LSYNROVPAGIGAMPALHTLNIDNIDLPAPAFQOLGSLDITNSNLTATL
PDPLESRGDADASPAPVLVSFGPLHNCBULMRLARPDLETASPPELAGRY
PAAVEGEFSCPEPIIARTQRLWLEGGATLRCALDPAFTMHWVGPDDLVGNS
SRAAPFNGTLEIGATGADAGATCTIATNPAGEATRVLRVLAIPHGNSAEGGR
PSPSIIAAGARPAEGEGLLESEPAVQTEVATSGLVSGRPAADPVMFOIYNS
SDRLITVIVPASGHHPLKLIVPGADVLDICLALSPAAGPSDLTATRLGCAHFT
LPASPLCALQAHVIGLITTAVGAVVALIVTVALLVIRGAGNGRLPLKLSHYO

```

SOJNGSPPTKAPHPSPRPPRQSCSLDLSGAGCYGVARLUGAMARSHVHGL  
 LGAGCGVGGASABRLBSYV"

Query Match 83.1%; Score 1585.4; DB 9; Length 1889;  
 Best Local Similarity 99.9%; Pred. No. 1.1e-210;  
 Matches 1586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 322 CTGGTGGAGCTGGGCAACCGGAGCGCTCCGGGGGCCCCCTCATATCTGAGAGCACTCATCTC 381  
 Db 1 CTGGTGGAGCTGGGCAACCGGAGCGCTCCGGGGGCCCCCTCATATCTGAGAGCACTCATCTC 60  
 QY 382 AGCGGCAACCAAGCTGGGCGCCATCGCGCGGAGCTTTCAGACGACTTCTAGAGAGCTG 441  
 Db 61 AGCGGCAACCAAGCTGGGCGCCATCGCGCGGAGCTTTCAGACGACTTCTAGAGAGCTG 120  
 QY 442 GAGGACTTGAACCTGTCTTCAACACCTCCGCGAGGTGCTTGGGCGGCGCATCGGCGC 501  
 Db 121 GAGGACTTGAACCTGTCTTCAACACCTCCGCGAGGTGCTTGGGCGGCGCATCGGCGC 180  
 QY 502 ATGCTGCGCCCTGACACCTCAACCTTGAACCTTATGACGAGCTGCGCGCGCAGC 561  
 Db 181 ATGCTGCGCCCTGACACCTCAACCTTGAACCTTATGACGAGCTGCGCGCGCAGC 240  
 QY 562 GCGTTGCGCCAGCTCGGTCAAGCTCTCCGCGCTGAGCTCACTCAACCGCTGAGCAG 621  
 Db 241 GCGTTGCGCCAGCTCGGTCAAGCTCTCCGCGCTGAGCTCACTCAACCGCTGAGCAG 300  
 QY 622 CTGGGCTTCGGAACCGGCTTTCTCTCGTGGGCGTGAATGACAGAGCTTCTCCGCGCTG 681  
 Db 301 CTGGGCTTCGGAACCGGCTTTCTCTCGTGGGCGTGAATGACAGAGCTTCTCCGCGCTG 360  
 QY 682 GTGCTAGCTTTTGAAGCGGAAACCCCTGACATGCACTGAGAGCTGAGCGTGGGCTGCGCG 741  
 Db 361 GTGCTAGCTTTTGAAGCGGAAACCCCTGACATGCACTGAGAGCTGAGCGTGGGCTGCGCG 420  
 QY 742 CTGGCGCGCGCGGACGACCTGGAACGTCGCTCCCGCGCGCGCTGAGCGCGCTAC 801  
 Db 421 CTGGCGCGCGCGGACGACCTGGAACGTCGCTCCCGCGCGCGCTGAGCGCGCTAC 480  
 QY 802 TTCTGGGCAATGCGCGGAGCGGCTTCTCTGAGAGCGCGCTCAATTGCGCGCGCACG 861  
 Db 481 TTCTGGGCAATGCGCGGAGCGGCTTCTCTGAGAGCGCGCTCAATTGCGCGCGCACG 540  
 QY 862 CAGCGGCTCTGGGCTGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 921  
 Db 541 CAGCGGCTCTGGGCTGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
 QY 922 CCGCGCGCTACCATGACATGAGTGGTCTGACGACCGGTTGGTGGCACTCTCCGA 981  
 Db 601 CCGCGCGCTACCATGACATGAGTGGTCTGACGACCGGTTGGTGGCACTCTCCGA 660  
 QY 982 GCGCGGCGCTTCCCAACCGGAGCTTGAAGATTGGGGTGAACCGGCGCTGGGAGCTGGG 1041  
 Db 661 GCGCGGCGCTTCCCAACCGGAGCTTGAAGATTGGGGTGAACCGGCGCTGGGAGCTGGG 720  
 QY 1042 GGTACACCTGATGCGCACCAACCTGCTGGTGAAGCGGCGGCGGCGGCGGCGGCGG 1101  
 Db 721 GGTACACCTGATGCGCACCAACCTGCTGGTGAAGCGGCGGCGGCGGCGGCGGCGG 780  
 QY 1102 GTGCTGGCTTGGCGGCGGAGAGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161  
 Db 781 GTGCTGGCTTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
 QY 1162 GACATGCGCGGCTCGGCTGCACTGTGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1221  
 Db 841 GACATGCGCGGCTCGGCTGCACTGTGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
 QY 1222 GCGGTGACAGGTGACGAGAGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1281  
 Db 901 GCGGTGACAGGTGACGAGAGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
 QY 1282 CCAAGCGACCCAGGTGATGTTCCAAATCGATCAACAGCAGGAGAGATGAGACCTTC 1341

Db 961 CCAAGCGACCCAGGTGATGTTCCAAATCGATCAACAGCAGGAGAGATGAGACCTTC 1020  
 QY 1342 ATTCACCGAATTTGCCAGGCTCCAGGCGACCACTTCTGTGTAAGAGACCTGTCCCGG 1401  
 Db 1021 ATTCACCGAATTTGCCAGGCTCCAGGCGACCACTTCTGTGTAAGAGACCTGTCCCGG 1080  
 QY 1402 GTTACTATGACCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1461  
 Db 1081 GTTACTATGACCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 QY 1462 GCGACGAGGCTGTGGGCTGTGCGCCATTTCACAGCTGCGCGGCTTCCGCTGTGCGAC 1521  
 Db 1141 GCGACGAGGCTGTGGGCTGTGCGCCATTTCACAGCTGCGCGGCTTCCGCTGTGCGAC 1200  
 QY 1522 GCGCTCAGGCGCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1581  
 Db 1201 GCGCTCAGGCGCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 1582 GCTGCTTACTGCTTCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1641  
 Db 1261 GCTGCTTACTGCTTCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
 QY 1642 CGCTCCCGCTCAAGCTCAGGCGACGTCAGTCCAGACCAATGAGAGCGCCAGCCAC 1701  
 Db 1321 CGCTCCCGCTCAAGCTCAGGCGACGTCAGTCCAGACCAATGAGAGCGCCAGCCAC 1380  
 QY 1702 CCGAAGGCCACCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1761  
 Db 1381 CCGAAGGCCACCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440  
 QY 1762 CTGGAGAGTGGCGGCTGTCTACGTTATGCGAGGCGCTTGGAGAGCTTGGCTCGGAC 1821  
 Db 1441 CTGGAGAGTGGCGGCTGTCTACGTTATGCGAGGCGCTTGGAGAGCTTGGCTCGGAC 1500  
 QY 1822 AGCACTCTGTGATGGGAGGCTGCTCGGAGGAGGCTGCGGAGGCTTAGAGGCGAGCGC 1881  
 Db 1501 AGCACTCTGTGATGGGAGGCTGCTCGGAGGAGGCTGCGGAGGCTTAGAGGCGAGCGC 1560  
 QY 1882 GAGCGGCTGGAAGAGAGTGTGTGA 1908  
 Db 1561 GAGCGGCTGGAAGAGAGTGTGTGA 1587

RESULT 5  
 BC023036 2553 bp mRNA linear ROD 30-JUN-2004  
 LOCUS  
 DEFINITION Mus musculus leucine rich repeat and fibronectin type III domain containing 4, mRNA (cdna clone MGC:36464 IMAGE:5358628), complete cds.  
 ACCESSION BC023036  
 VERSION BC023036.1 GI:27696757  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2553)  
 Strassburg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, K.H., Shemmer, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marnett, K., Farmer, A.A., Rubin, G.M., Hong, L., Stadelman, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tothilyuk, S., Carninci, P., Prange, C., Rana, S.S., Loguigliano, N.A., Peters, G.J., Abramson, R.D., Mollath, S.J., Bosak, S.A., McMan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,



Db 1349 AGCGGCGCTGAGATGAGAGAGCCTATATCTGATTCGACCAACCTCTGTGTAGAGCC 1408

Qy 1081 ACAGCCCGAGTAAAGTACGCGGCTGTGCTGCTTGGCCCATGTTGGGAACAGAGGCCGAG 1140

Db 1409 ACAGCCCGAGTAAAGTACGCGGCTGTGCTGCTTGGCCCATGTTGGGAACAGAGGCCGAG 1468

Qy 1141 GGG 1200

Db 1469 GGG 1528

Qy 1201 GGGAGCGCTGAGATGAG 1260

Db 1529 GGAACCTTAAATATGAAACAG 1588

Qy 1261 GTGAGCTGAGGCTCCCGGG 1320

Db 1589 GTGAGCTGAGGCTCCCGGG 1648

Qy 1321 AGCAGCGAATGAG 1380

Db 1649 AGCAGCGAATGAG 1708

Qy 1381 CTGAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440

Db 1709 CTGAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1768

Qy 1441 GCTGAGGCTCTGAGCTCAGGCGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

Db 1769 GCTGAGGCTCTGAGCTCAGGCGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1828

Qy 1501 CCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

Db 1829 CCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1888

Qy 1561 GCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620

Db 1889 GCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1948

Qy 1621 GCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

Db 1949 GCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2008

Qy 1681 AATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

Db 2009 AATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2068

Qy 1741 CAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797

Db 2069 CAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2128

Qy 1798 CTGGAGAGAGCTTGGGCGCGAGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1857

Db 2129 CTGGAGAGAGCTTGGGCGCGAGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2188

Qy 1858 TGCCGGGGGGGTGAG 1908

Db 2189 TGCCGGGGGGGTGAG 2239

RESULT 6  
BC023156 2612 bp mRNA linear ROD 25-VAR-2004  
LOCUS Mus musculus leucine rich repeat and fibronectin type III domain  
DEFINITION containing 4, mRNA (cDNA clone MGC:36545 IMAGE:4951386), complete  
cds.  
ACCESSION BC023156  
VERSION BC023156.1 GI:22137562  
KEYWORDS MGC  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1, (bases 1 to 2612)

# AUTHORS PUBLISHED REFERENCE AUTHORS TITLE JOURNAL

## REMARK COMMENT

## FEATURES source

## gene

## CDS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauener, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buettow, R.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uddin, T.B., Roshyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.U., Abramson, R.D., Mulich, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wierley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schultz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932  
2 (bases 1 to 2612)  
Strausberg, R.  
Direct Submission  
Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H., Kowis, C.R., Sneed, A.U., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
Series: IRAP Plate: 60 Row: K Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23510244.  
Location/Qualifiers  
1..2612  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:36545 IMAGE:4951386"  
/tissue\_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."  
/clone\_lib="NCI CGAP Mam"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..2612  
/gene="Lrln4"  
/note="synonym: MGC36545, MGC36464"  
/db\_xref="locusID:225875"  
/db\_xref="MGI:2385612"  
372..2282  
/gene="Lrln4"  
/codon\_start=1  
/product="Lrln4 protein"  
/protein\_id="AAH23156.1"  
/db\_xref="GI:22137565"  
/db\_xref="locusID:225875"



/db\_xref="MGI:2385612"  
translation="MAPLLLLLLASGAACPDPVCONLSSESLTICARGLLFPVP  
NDRRTVELADNPDLGPDPRNMGTGLDILTSRAIRIRISAGHJEDLSRSH  
LDGRNLVLESSSIRGPNVHLLISGQRIARAGADDDLDLSIEDLDVRYNNLROY  
FWAGISMPALHTLMDLHNLIDALPVGFAQLSDLSRLDLSNLTATLAPLRSRGR  
DAEASPSPLVLSFSGNPLHNCCELLMLRLARPDLLETGASPLAFYFAVEGER  
SCEPPLIARHTORLWLEBGRATLRCALGDPVPMHVGPDRLVGSSSPAMAFPNG  
TLEIGTCAGAGAYTCIATNPAGBATRVELRYVALPHGNTSAEGSRPSPDLIAS  
ARTAAEGGTLESBPAYQVTEVTATSGIVSGLGRPADPVMFPOIYVNSSEDTLIYR  
IVPASSHHFLKHYRPGADYDPLCLIALSPAGSPDLTRLLGCARHSTLEPANLCHA  
LOAHVGGTLTVAAGVLAALVFTVALTVRGKAGNRLPLKLSHVQSTNGSTIS  
MPKSHPSPPRPQRCSLDLGDITGCGYARLRGAMARSHSVHGLGACRGV  
GSAERLESYV"  
1071..1208  
/gene="Lrtn4"  
/note="LRCT; Region: Leucine rich repeat C-terminal  
domain"  
/db\_xref="CD:smart00082"  
1209..1472  
/gene="Lrtn4"  
/note="Tgcam; Region: Immunoglobulin domain cell adhesion  
molecule (cam) subfamily"  
/db\_xref="CD:cd00931"  
1596..1805  
/gene="Lrtn4"  
/note="Fn3; Region: Fibronectin type III domain"  
/db\_xref="CD:pfam00041"

misc\_feature  
79.9%; Score 1523.8; DB 10; Length 2612;  
Best Local Similarity 87.7%; Pred. No. 3.6e-202;  
Matches 1676; Conservative 0; Mismatches 232; Indels 3; Gaps 1;

ORIGIN

Query Match 79.9%; Score 1523.8; DB 10; Length 2612;  
Best Local Similarity 87.7%; Pred. No. 3.6e-202;  
Matches 1676; Conservative 0; Mismatches 232; Indels 3; Gaps 1;

1 ATGCCCCCGCGCTCTGCTGCTGCTGCGCACTGAGCGCGCCCTGCTGCC 60  
372 ATGCCCCCGCGCTCTGCTGCTGCTGCGCACTGAGCGCGCTGCTGCC 431  
61 TGGCTGCGCAGAACCTGCGAGTCCGCTCAGACCCCTCTGCGCCACCGAGCTGCTG 120  
432 TGTGTGCGCAGAACCTGCGAGTCCGCTCAGACCCCTCTGCGCCACCGAGCTGCTG 491  
121 TTTGTGCGCGCCCAAGCGTGAGCGCGCAGAGTGAAGCTGCGCTGCTGACAACTTCATC 180  
492 TTTGTGCGCAGAACCTGCGAGTCCGCTCAGACCCCTCTGCGCCACCGAGCTGCTG 551  
181 CAGGCGCTGCGCGCCCTGCTGCTGCGCACTGAGCGGACTGCTGCTGCTGCTGCTGCT 240  
552 CAGGCGCTGCGCGCCCTGCTGCTGCGCACTGAGCGGACTGCTGCTGCTGCTGCTGCT 611  
241 CGCAATGCGCATACCGGCACTTGGGCGCGCGCTTGGGGAAGCTGAGAGCTGCGTTC 300  
612 CGAAATGCGCATACCGGCACTTGGGCGCGCGCTTGGGGAAGCTGAGAGCTGCGTTC 671  
301 CTCCACCTTGACGGCAACAGGCTGTGAGCTGGGCGACCGGAGCTTCGCGGCGCTGTC 360  
672 TTGCACTCGATGGCAACAGGCTGTGAGCTGGGCGACCGGAGCTTCGCGGCGCTGTC 731  
361 AATTCGACGACTCTCTCAGCGGCAACAGCTGGGCGCGCATGCGCGCGGAGGCTTC 420  
732 AACCTCCAGCATCTCATTTCTCAGTGGCAATCAACTGGCGCGCATGCGCGCGGAGCTTC 791  
421 GAGCACTTCTGAGAGCGCTGAGGAGCTGAGCTTCTCTCAACAACAACCTCGCGAGGTG 480  
792 GAGCACTTCTCTTGACAGTCTCGAGAGCTGAGTGTCTCTATTAACAACCTCGCGAGGTG 851  
481 CCTTGCGCGCGCATGCGCGCATGCTGCTGCTGCGACACCTCTCAACTGAGCACTTCT 540  
852 CCTTGCGCGCGCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911  
541 ATTGAAGCACTGCGCGCGCGCTGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
912 ATGAGAGCACTACCCCGAGGTGCTTGTGCTGAGCTGAGCGAGCTGCTGCTGCTGCTGCTG 971

601 ACTCCAAACCGCTGAGCGACGCTGCTGCGGACCGCTTTTCTCTGCTGCGCTGATCA 660  
972 ACTCCAAACCGCTGAGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1031  
661 GAGGCTCTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
1032 GAGGCTCTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1091  
721 GAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
1092 GAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1151  
781 CCGGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
1152 CCAACCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1211  
841 CCGCTGATTTGCGCGCGCACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
1212 CCGCTGATTTGCGCGCGCACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1271  
901 CCGCTGATTTGCGCGCGCACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
1272 CCGCTGATTTGCGCGCGCACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1331  
961 TTTGTTGGCAACTCTCTCCGAGCGCGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
1332 CTGTTGTCACACTCTTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1391  
1021 ACCGCGCTGCGGAGCGCTGCGGCGCTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
1392 ACGGCGCTGCGGAGCGCTGCGGCGCTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1451  
1081 ACAGCGCGAGTAGAATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
1452 ACAGCGCGAGTAGAATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1511  
1141 GGGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
1512 GGGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1571  
1201 GGGAGCGCTGAGCTCTGAGCTGAGCGCTGAGCTGAGCGAGTGAACCGGCACTCAGGGCTG 1260  
1572 GGAATTTAGATTTGAACCAAGCACTGAGCGAGTGAACCGGCACTCAGGGCTG 1631  
1261 GTGAGCTGGGCTGCGGCGCGCGCGAGCCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1320  
1632 GTGAGCTGGGCTGCGGCGCGCGCGAGCCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1691  
1321 AGCAGCGAAGATGAGACCTCTATCTTACCGGATTTGCCAGCTCCAGCCACCACTTCTG 1380  
1692 AGCAGCGAAGATGAGACCTCTATCTTACCGGATTTGCCAGCTCCAGCCACCACTTCTG 1751  
1381 CTGAAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
1752 CTGAAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1811  
1441 GCTGAGCGCTGAGCTGAGCGCGCGCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1500  
1812 GCTGAGCGCTGAGCTGAGCGCGCGCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1871  
1501 CCGGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
1872 CCGGCGCGCTGAGCTGAGCGCGCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1931  
1561 GCGCTGAGCGCGCTGAGCTGAGCGCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1620  
1932 GCGCTGAGCGCGCTGAGCTGAGCGCGAGCGAGTGGATTTCCAAATCCAGTAAAC 1991  
1621 GCGCGCGCGCGCGAGTGGATTTCCAAATCCAGTAAACCGTTCAGTTCAGTTCAGTTCAG 1680  
1992 GCGCGCGCGCGCGAGTGGATTTCCAAATCCAGTAAACCGTTCAGTTCAGTTCAGTTCAG 2051  
1681 AATGAGCGCGCGCGAGTGGATTTCCAAATCCAGTAAACCGTTCAGTTCAGTTCAGTTCAG 1740



Db	132791	AGCAGCGAAGATGAGACCCCTCATCTACCGGTGAGGATGCGTCCCAACCCCGCTGCAC	132850
QY	1349	-----	1348
Db	132851	TCCCGAGGCGCTTCCCTCCGACCTCTGCTCCCAACAGGCTTTGTTCCACCCCTCTCTC	132910
QY	1349	-----	1348
Db	132911	TCGAGGAGTCACACCCCTCCCGACCATGAGCTGCTGCACTCTTGAGAGAGCATGAGCCT	132970
QY	1349	-----	1348
Db	132971	GGCCTGAGCTTGCTGCTGAGAGGAGGCTGAGAGCTCAAGGAGCGGACAGACCCACCC	133030
QY	1349	-----	1348
Db	133031	GAGATGTCCTCCAGGTGTCACTGAGCTGTGACAGGCTCCCGGACAGAGTGGCGATTGGGT	133090
QY	1349	-----	1348
Db	133091	CTGTCTGAGACATAGGCGCGCGCGCGGTGTATTATTCTGCTCTGCTGTGCATCCCC	133150
QY	1349	-----	1348
Db	133151	CCACCCCATTCCTCCCTGCTGCTGCGTCTGTCTGTGCGAGTGCACACACCATCTCCTGGC	133210
QY	1349	-----	1348
Db	133211	ACACTCTTCAGCCTGTTTCACTTCTTAGTGAACCCAGAGCCCGAGTGGCCCCAGGGGG	133270
QY	1349	-----	1348
Db	133271	AGGAGTGTGTTGAGCTGGGAGCAGGAGTGGGGAGGTGAGGATGAGGCTTAGACCCAGAC	133330
QY	1349	-----GGATTGTCCAGCCTCCAGCCACCATCTTC	1337
Db	133331	CCCGGCTCTGACCAACCTGACCCCTTGCTGAGAGATTGTGCCAGCTCCAGGCCACCATTC	133390
QY	1378	CTGCTGAGACACTGTCCTCCCGGCGCTGACTTATGACTCTGCTGCTGGCTTGTCAACG	1437
Db	133331	CTGCTGAGACACTGTCCTCCCGGCGCTGACTTATGACTCTGCTGCTGGCTTGTCAACG	133450
QY	1438	GCCGCTGGGCGCTCTGACCTTCAAGGCCACAGGCTGTGGGCTGTGCCATTCTTCACG	1497
Db	133451	GCCGCTGGGCGCTCTGACCTTCAAGGCCACAGGCTGTGGGCTGTGCCATTCTTCACG	133510
QY	1498	CTGCGCGGCTCGGCGCTGTGGCAGGCGCTGAGAGGCCAAGTGTGGGCGGAGCCCTGAC	1557
Db	133511	CTGCGCGGCTCGGCGCTGTGGCAGGCGCTGAGAGGCCAAGTGTGGGCGGAGCCCTGAC	133570
QY	1558	GTGCGCGTGGGGGTGTGCTGTGTGCTGCTGCTTACTGTCTTCACTGTGGCTTGTGATT	1617
Db	133571	GTGCGCGTGGGGGTGTGCTGTGTGCTGCTGCTTACTGTGTCTTCACTGTGGCTTGTGATT	133630
QY	1618	CGGGGCGGGGGGGCGGAAATGGCGGCTTCCCTCAAGCTTACAGCCATGCTCCAGTCCAG	1677
Db	133631	CGGGGCGGGGGGGCGGAAATGGCGGCTTCCCTCAAGCTTACAGCCATGCTCCAGTCCAG	133690
QY	1678	ACCAATGAGGCGCCCAAGCCCAACCCAGGCGCCACCGCGGAGCCCGCGCCCGG	1737
Db	133631	ACCAATGAGGCGCCCAAGCCCAACCCAGGCGCCACCGCGGAGCCCGCGCCCGG	133750
QY	1738	CCGACGCGCAGCTGCTCTTCTGGAAGCTTGGGAGATGCCGGGTGTCTACGTTATGCAAGCGC	1797
Db	133751	CCGACGCGCAGCTGCTCTTCTGGAAGCTTGGGAGATGCCGGGTGTCTACGTTATGCAAGCGC	133810
QY	1798	CTGGAGAGAGCTTTGGGCGCCGACGGAAGCCACTCTGTGCTATGGGAGGCGTGTCTGGGCGAGGG	1857
Db	133811	CTGGAGAGAGCTTTGGGCGCCGACGGAAGCCACTCTGTGCTATGGGAGGCGTGTCTGGGCGAGGG	133870
QY	1858	TGCGCGGAGGTGAGAGGACGCGCCGAGCGGCTTGGAAAGAGTGTGTGTGA 1908	

[illegible]





QY 721 GAGTGTGTGTGCTGCGGCGGCTGAGCGGCGCGGACGACCTGAAAAGTGTGCTCCCG 780  
 Db 5047 GAGTGTGTGTGCTGCGGCGGCTGAGCGGCGCGGACGACCTGAAAAGTGTGCTCCCG 4988  
 QY 781 CCGGCGCTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 840  
 Db 4987 CCGGCGCTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4928  
 QY 841 CCCCTCATTTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 900  
 Db 4927 CCCCTCATTTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4868  
 QY 901 CCGTGTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 960  
 Db 4867 CCGTGTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4808  
 QY 961 TTGGTTGGCAACTCTCTCCGAGCGGCTTTCCCGACGCGGACCTTAGAGATTGGGGTG 1020  
 Db 4807 TTGGTTGGCAACTCTCTCCGAGCGGCTTTCCCGACGCGGACCTTAGAGATTGGGGTG 4748  
 QY 1021 ACCGCGCTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1080  
 Db 4747 ACCGCGCTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4688  
 QY 1081 ACAGCCCGAGTAGAACTGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1140  
 Db 4687 ACAGCCCGAGTAGAACTGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4628  
 QY 1141 GGGGGCGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1200  
 Db 4627 GGGGGCGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4569  
 QY 1201 GGGAGCGCTGAGTAGAACTGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1260  
 Db 4568 GGGAGCGCTGAGTAGAACTGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4509  
 QY 1261 GTAGAGCTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1320  
 Db 4508 GTAGAGCTGCGGCGGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4449  
 QY 1321 AGAGAGGAGATAGAACCTTATCTACCGGATTTGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1380  
 Db 4448 AGAGAGGAGATAGAACCTTATCTACCGGATTTGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4389  
 QY 1381 CTGAGACGCTGCTCTCCCGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1440  
 Db 4388 CCGGCGATCTCTCTCCCGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4329  
 QY 1441 GCTGGCGCTCTGAGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 1482  
 Db 4328 CTTGGGCGCTGAGCGGCGGCTGAGCGGCGGACGACCTGAAAAGTGTGCTCCCG 4287

RESULT 10  
 BC007718 1844 bp mRNA linear PRI 12-FEB-2004  
 LOCUS Homo sapiens leucine rich repeat and fibronectin type III domain  
 DEFINITION  
 ACCESSION BC007718  
 VERSION BC007718.2 GI:33873739  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS  
 1 (bases 1 to 1844)  
 Strusberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,L., Shemmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

REMARK  
 COMMENT  
 JOURNAL  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Carninci,P., Prange,C., Raha,S.S., Loggiano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahy,J., Helton,E., Ketterman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,D., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,  
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1844)  
 Strusberg,R.  
 Direct Submission  
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:14712783.  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
 Hansen,N., Ho,S.-L., Karlins,E., Kong,P., Laric,P., Legapi,R.,  
 Maduro,Q.L., Masilio,C., Maskell,B., Mastrian,S.D., McLooney,J.C.,  
 McDowell,J., Pearson,R., Stantipop,S., Thomas,P.J., Touchman,J.W.,  
 Tsugeon,C., Vogt,J.D., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
 Young,A., Zhang,L.-H. and Green,E.D.

FEATURES  
 source  
 1.1844  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:410922"  
 /tissue\_type="Muscle, rhabdomyosarcoma"  
 /clone\_lib="NIH MGC 17"  
 /lab\_host="DH10B-R"  
 /note="vector: POTB7"  
 1..>1844  
 /gene="LRFN4"  
 /note="synonym: MGC3103"  
 /db\_xref="LOCUSID:78999"  
 221..>1844  
 /gene="LRFN4"  
 /codon\_start=1  
 /product="LRFN4 protein"  
 /protein\_id="AAH07718.1"  
 /db\_xref="GI:33873740"  
 /db\_xref="LOCUSID:78999"  
 /translation="AAPLLILLASGAACPDPVCONLSSTLCAHRLGLVPP  
 NVDRRTVELRADNFIQALGPDPFNNMGLVDTLSRAVIRIGARGFDSLSLH  
 LDGNLVEIGTGSLEGPVNLQHLISGQDLGRIPGADPDFLESLDLSYNNLRV  
 PWAGIGAMPALHTNLIDNLIDLPAPAFQAGQGRSDLSNRLATLADPLFSGR

gene  
 CDS











Db	212631	AAGAGCCACCCACCAAGGAGCCCTCCACCCCTGCCAGCGAGCTGCCCTGACCTTA	21269
QY	1765	GGAGATGCC---GGTGCTACCGGTTATGCGACGGCGCTTGAGAGAGCTTTGGCCCGACGG	1821
Db	212691	GGAGACACCGCGTGAGATGCTACCGGATATGCGACGGCGCTTGAGAGAGCTTGAGCCCGCGCG	212750
QY	1822	AGCCACTGCTGTCATAGGGGGGGCTGCTCGGGGGAGAGGGTGTCCGGGGGGGTATAGAGGACACGGCC	1881
Db	212751	AGCCACTGTCATAGGGGGGGGGCTGCTCGGGAGAGTGGGTGTCCGGGGGTATAGGGGGAGATGGCG	212810
QY	1882	GAGCGGCTGGAAGAGAGTGTGCTGTGA	1908
Db	212811	GAGCGGCTGAGAGAGAGCTGTGTGTA	212837
RESULT 13			
	AX358338	1653 bp	DNA
LOCUS	AX358338		linear
DEFINITION	Sequence 9 from Patent WO0202604.		PAT 13-FEB-2002
ACCESSION	AX358338		
VERSION	AX358338.1	GI:18674974	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1		
TITLE	Thayer,E.C., Sheppard,P.O. and Presnell,S.R.		
JOURNAL	Leucine-rich repeat proteins, zlr7, zlr8 and zlr9		
FEATRES	Patent: WO 0202604-A, 9 10-JAN-2002;		
source	ZymoGenetics, Inc. (US)		
	Location/Qualifiers		
	1..1653		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
ORIGIN	/note="Degenerate polynucleotide sequence"		
Query Match	57.7%; Score 1101.8; DB 6; Length 1653;		
Best Local Similarity	55.7%; Pred. No. 1.4e-143;		
Matches	920; Conservative 349; Mismatches 381; Indels 2; Gaps 2;		
QY	1	ATGCCCCCGCGCTCCTGCTGCTGCTGCGCAGTGAAGCGGCGCTGCCCGCTGCC	60
Db	1	ATGCGNCCNCCNYTNYTNYTNYTNYTNGCNMWSNGNCGNCGNCGNCGNCGN	60
QY	61	TGGCGCTGCGAAGACTGTCGAGTGCCTCAGACCCCTCTGCGCCACGAGGCGCTG	120
Db	61	TGGYNTGTGCAAPAAAYTMSNAGRWMSNYTMSNACNYTGTGCAAMGNGNYTNYN	120
QY	121	TTTGTGCGCGCCACAGCGTGAACCGCGCGCAGTGAAGCTCGGCTGCTGCAACTTCATC	180
Db	121	TTTGTGNCNCCNAAAGTNGAAYMGNMNAONGTNGAAYTMMGNTTNGCNAAYATYATH	180
QY	181	CAGGCGCTTGGGGCCCCCTACTTCGCGCAATGACGGGACTGTGTGACACTGTCT	240
Db	181	CARGNNTYNGNCCNCGAAYTTYMNAAYATGAONGNYTNGTNGAAYYTNA.CNYTMSN	240
QY	241	CGCAATGCGATCAACCGCATTTGGGGCGCGGCTTTGGGGAGCTCGAGAGCTGCTCC	300
Db	241	MGNAAVGCNAATHACMNGNAITHGNG.CNMNGCNTTGTGGNGAAYTNGAARMSNYTMMGWSN	300
QY	301	CTCCACCTTGACGGGCAAGCGCTGTTGAGAGCTGGGACCGGGAGCTCCGGGGCGCGTC	360
Db	301	YTNCAAYTNGAAGNNAAYMGTNTNGARITYTNGNA.CNCGMWSNYTMMGANGCNGTN	360
QY	361	AATCTGACGACACTCTACTCTCAGCGGCAACGAGCTGGCGCGCATCGCGCGGAGCCTTC	420
Db	361	AAYTNCARCAAYTNAATHYTMSNGNAAVCAARYTNGMNGMNAITHGNCNCGANGCNTTY	420
QY	421	GACGACTTCTAGAGAGGCTGAGAGACTGTGACTGTCTTACAACAACACTCGGCGAGTG	480
Db	421	GAYGAYTYTNTGARSNTYNGARGYATYTNGAAYTWTMSNTAAYAAVAYTMMNCARGTN	480

## ORIGIN

Query Match	57.7%;	Score 1101.8;	DB 6;	Length 1653;
Best Local Similarity	55.7%;	Pred. No. 1.4e-143;		
Matches 920;	Conservative 349;	Mismatches 381;	Indels 2;	Gaps 2;

Qy	481	CCCTGGGCGCGGCAATCGGCGGCAAGCCGCGCCCTGGACACCCGCAACCTGGACCAATTAACCT	540
Db	481	CCNTGGCGCGGATHHGGNCAATGCCGCGCNYTNCATACNYTAAAYTMMGACATAAAAYTN	540
Qy	541	ATTGACCGCACTGCGCCCGAGCGGCGCTTGGCCAGCTGGGTGAGCTCTCCGCGCTTGACCTC	600
Db	541	ATHHAYGCAANTTCCGCGCGGNGGCTTGTGCMCARTYTNGNRCARITYNMNMGNITYNGAAYTN	600
Qy	601	ACCTCCAACCGGCTGGCCACGCTGGCTCCGAGCCGCTTTTCTCTGCTGGGCGGTGATGCA	660
Db	601	ACNMNSAAAYMGNTYTMGNCANITYTGCNCNGAACNANTYNTTYSMMGMGNGMGAGCGN	660
Qy	661	GAGGCGCTCTCCCGCGCGCGCTGGTGAAGCTTAAAGGCGGAAACCCCGTGCACCTGCAACTGT	720
Db	661	GARGCMMNSCCNGNCCNANTYNTGNTYNTWMSYTMSSNGNAAAYCCNITYNCAITYGAAAYTG	720
Qy	721	GAGCTGCTGTGGCTGCGGCGGCTGGCGCGGCGCGGACCACTGGAAAGTGCGGCTCCCGG	780
Db	721	GARTYNTTNTGGYTTMMGMGNITYTNGCMGMGNCNGAYAYTNGABACNTGYGCMNSCCN	780
Qy	781	CCCGGCGCTGGCGCGGCGGCTACTTCTGGGCGAGTGCCCGAGGGCGAAGTTCTCTGTGAGCCG	840
Db	781	CCMGANTYTMGCGNMGNMNTAYTTYTGCGCMGTNCNMGARGNGARITYWMSNTGYGACCN	840
Qy	841	CCCGCTCATTTGGCGGCGCACACGACGCGCTGTGGGGTGTTGAGAGGCGCGGCGGACCGCTG	900
Db	841	CCNTYNTATHTTGCMNMCAYACNCAKMGNTYTTGGTNTYNGARSGNCARMGNGCNACTYN	900
Qy	901	CGGTGGCGGCGGCGCTGGGGTGAACCCCGCGGCTACCATGCACTGGGTGGTCTCTGACGACGG	960
Db	901	MGNTGYMGNGCNYTNGNGAYCCNGGCMCNACNATGCAITYGGGTNGGNCNGAYGAYVMGN	960
Qy	961	TTGGTTGGCAACTCTCTCCCGAGCCCGGGGCTTTCCCGCAACGGGACCTTAGAGATTGGGGTG	1020
Db	961	YTNGTGTGNAAYWMSNMGMNGCMGNGCNTTYCCNAAAYGNAANYTNGARATHGGAGTN	1020
Qy	1021	ACCGGCGGTGGGAGCGCTGGGGGGCTACACCTGATCGGCAACAACTGTGCTGGTGAAGCC	1080
Db	1021	ACNGSNGCNGSNGAYCNGSNGSNTAYACNTGYATHGCMACNAAAYCCNGCNGSNGAGCN	1080
Qy	1081	ACAGCCCGAGTAGAAGCTCGGGGTGCTGAGCCTTGCCCCATGGTGGGAAACAGACGTGCCAG	1140
Db	1081	ACNGCMNGNGTNGARYTNMGNGTNTYTMGCMYTNCCNCAAYGNGGNAAYWMSMSNGCNGAR	1140
Qy	1141	GGGGGCGCGCCCGGGGCGCTGGGACATGCGCGGCTCCGCGCTGGCAATGCTGCCGAGGGTGAG	1200
Db	1141	GGNGGNGNCCNGGNCNWSNGAYATHGCMGMSNCCNMGNACNGCNGCNGAGGNGAR	1200
Qy	1201	GGGAGCGCTGAGCTGTGAGCAGCGCGGTGAGGGTGAAGGAGTGAACCGCACCTCGAGGCGTG	1260
Db	1201	GGNACNTYTNARWSNGARCCNGCNGTMCARGTNAENARGTNACNGNACNWSMGNTYN	1260
Qy	1261	GTGAGCTGGGGTCCCGGGCGGCGACGCGGACCCAGTGTGATGTTCCAAATCCAGTACAC	1320
Db	1261	GTMSNTTGGGNCNCGNGMGNCNGCNGAAYCCMGNTTGAATGTTTCARATHCARATAAAY	1320
Qy	1321	AGCAGCGCAAGATGAGACCCCTCATCTACCGGATTTGCCACGCTCCACGCAACCACTTCCGT	1380
Db	1321	WSNMSNGARAYGARACNYTAAHTATMGNAHTGTNCNCSNMSNWCAYCAATTYYTN	1380
Qy	1381	CTGAAGCACTCGTCCCGGCGGCTGACATAGCACTGTGCGTGGCGGCTTGTGCAACGCGGC	1440
Db	1381	YTNAARCAAYTNGTNCNCGSNGCNGATYATGAYTYTNTGYTNTYNGCNYTNMWSNCCNGCN	1440
Qy	1441	GCTGGGCGCTGTGACTTCACGGGCAACGAGGTGTGGGCTGTGGCCATTTCTCCACGCGTG	1500
Db	1441	GCNGGNCNMSNGAYTYTNAACNACNMGNITYTNTGNTGYGNCNCAITYTYSNACNYTN	1500
Qy	1501	CGGCGCTTGGCGCGCTGTGCACCGCCCTGACGAGCCGACGCTGTGGCGGCGGACCGTGAACG	1560
Db	1501	CCNGGMSNCCNANTYNTGYCAAGCNYTNCARCGNCAVGTNTYNGSNGNACNYTNACNGTN	1560

Ox	1561	GCCGGCGGCGGGGATGCCTGTGGTGGCCTTCATTGCTTTACTACATGG-CTTGCTGCTGCTC 1618
Dd	1561	GCNCTNGCNAGNCNTYTNCTNGCNCNTNAACNAGNYTNCAAYTGCGCNCNTGTGTGTTY 1620
Ox	1619	GGGGCGGGGGGGCGCGAAATGCGGCCCTCCCC 1650
Dd	1621	GGNCGNCGNCGNCCNCGAATGCGCNCNNNSCC 1652
RESULT 14		
LOCUS	AC106165	
DEFINITION	AC106165 269255 bp DNA linear HTG 13-MAY-2003 Rattus norvegicus clone CH230-102H22, WORKING DRAFT SEQUENCE, 3 unordered pieces.	
ACCESSION	AC106165	
VERSION	AC106165.4 GI:30578537	
KEYWORDS	HTG; HTGS_PHASEI; HTGS_DRAFT; HTGS_FULLTOP. Rattus norvegicus (Norway rat)	
SOURCE	Rattus norvegicus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (pages 1 to 269255)	
AUTHORS	Murzyn,D.,Maite, Metzker,M.,Lee, Abramzon,S., Adams,C., Alder,U., Allen,C., Allen,H., Alsbrooks,S., Amtn,A., Angiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnesrad,M., Benahmed,F., Blawid,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burck,P., Butrell,K., Calderon,E., Cardenas,J., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrick,D., Degado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Diyya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Dublin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Fall,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Gante,R., Garcia,A., Garner,T., Gazda,M., Gebregregys,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W., Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladin,S.L., Hodgson,A., Hogues,M., Holins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Joliviet,A., Karkachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C., Kowitz,C., Kraft,C.L., Lebow,H., Levay,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,Z., Lorenshewa,L., Lounseged,H., Lozada,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minga,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Muidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Mnakelemeh,O., Okunonu,G., Olarundusunon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Pioppert,F., Polindexter,A., Popovic,D., Pirmus,E., Pu,L.-L., Pizzo,M., Quinzio,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Rem,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,U., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Speed,A., Sodergren,E., Song,X.-Z., Sovelle,R., Soza,U., Steinle,M., Strong,R., Sutton,A., Swatke,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Thingle,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., White,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,K., Wiley,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yanub,S., Yen,J., Yoon,L., Yoon,V., Xu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zuo,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G., and Gibbs,R.A.	
TITLE	Direct Submission	

```

JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 269255)
AUTHORS      Worley,X.C.
TITLE        Direct Submission
JOURNAL      Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              3 (bases 1 to 269255)
REFERENCE    Rat Genome Sequencing Consortium.
              Direct Submission
              Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              On May 13, 2003 this sequence version replaced gi:23121847.
COMMENT       The sequence in this assembly is a combination of BAC based reads
              and whole genome shotgun sequencing reads assembled using Atlas
              (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
              in the feature table below represents a scaffold in the Atlas
              assembly (a 'contig-scaffold'). Within each contig-scaffold,
              individual sequence contigs are ordered and oriented, and separated
              by sized gaps filled with Ns to the estimated size. The sequence
              may extend beyond the ends of the clone and there may be sequence
              contigs within a contig-scaffold that consist entirely of whole
              genome shotgun sequence reads. Both end sequences and whole genome
              shotgun sequence only contigs will be indicated in the feature
              table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GJKL
Center clone name: CH230-102H22
----- Summary Statistics -----
Assembly program: Atlas 3.0;
Consensus quality: 240505 bases at least Q40
Consensus quality: 242683 bases at least Q30
Consensus quality: 244329 bases at least Q20
Estimated insert size: 251974; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

-- NOTE: Estimated insert size may differ from sequence length
-- (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
-- NOTE: This sequence may represent more than one clone.
-- NOTE: This is a 'working draft' sequence. It currently
-- consists of 3 contigs. The true order of the pieces
-- is not known and their order in this sequence record is
-- arbitrary. Gaps between the contigs are represented as
-- runs of N, but the exact sizes of the gaps are unknown.
-- This record will be updated with the finished sequence
-- as soon as it is available and the accession number will
-- be preserved.
FEATURES             source
     .
     .   1. 266736: contig of 266736 bp in length
     .   2. 266836: gap of unknown length
     .   3. 266837: contig of 1106 bp in length
     .   4. 267942: gap of unknown length
     .   5. 268043: contig of 1213 bp in length.
     .   Location/Qualifiers
         1. 269255
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-102H22"
            1. 1418
               /note="wgs_contig"
misc_feature
     1. misc_feature
Query Match          56.8%; Score 1083.8; DB 2; Length 269255;
Best Local Similarity 86.5%; Pred. No. 8,9e-142;
Matches 1196; Conservative 0; Mismatches 187; Indels 0; Gaps 0;
1 ATGAGCCCGGGCGGTCCCTGTCGTGCATGCAAGCAGTGAAGAAGGCCGCTGCGCGTGACC 60

```



Db	30853	ATGGCCCCGGCGCTCTATTGCTGTGTGTGTGGCCAGTGGAGGCGCCTGTGGCCGCTGCCC	30712
Qy	61	TGCGTCTGGCCAGAACTGTCCGAGTGCCTCAGCAACCTCTGTGTCCACCGAGGCTGTG	120
Db	30713	TGTGTGTGGCCAGAACTGTGTGGAGTGCCTCAGCAACCTTGTGTGCCACCGAGGCTGTGCTG	30772
Qy	121	TTTGTGGCGCCGCAACGTGTGACCGGGCGACAGTGTGGATGTGGTGGCTGATCAAACTTCAATC	180
Db	30773	TTTGTGTGCCAACGTGTGACGAGGCGACAGTGTGAATGTGGCTGTGCTGATCAAACTTCAATC	30822
Qy	181	CAGGCGCTTGGGGCCCCCTGTACTTCCGCAATGACGCGAATGTGTGACCTGACACTGTCT	240
Db	30833	CAGGCGCTTGGGGACCACTGTACTTCCGCAATGACAGGGTGTGTGACCTTAACATTTGTCT	30892
Qy	241	CGCAATGTCATCACCCGCATTTGGGGGCTCGGCGCTTTTGGGGACTCTGAGAGCTTGCCTTC	300
Db	30893	CGAATATGCATCACCCGCATTTGGGGGCTCGGCGCTTTTGGGGACTCTGAGAGCTTGCCTTC	30952
Qy	301	CTCCACCTTGTACGGGCACAGAGCTGTGTGAGACTGTGGGACCGGGAGCCTCCGGGGGCCCCGTC	360
Db	30953	TTTGACCTTGTATGGGCACAGAGCTGTGTGAGACTGTGGGACCGAGACGACCTTAAGGGGGCTGTCTC	31012
Qy	361	AATCTGCAGACCTCATCTCTCAGCGGCAACACAGCTGAGCGCGCATGTGGCGCCGAGCTTTC	420
Db	31013	AACTCTCAGACCTCATCTTCTTATGTGGCAATCAGCTGTGGCGCGAATCGCTCTGTGGGCGCTTTC	31072
Qy	421	GAGCACTTCTAGAGACCTGTGAGAGACCTGTGACCTGTCTTACAAACAACCTCGGACAGGTG	480
Db	31073	GAGCACTTCTTGTGACATCTTGTGAGAGACCTGTGAGTGTCTTATTAACAATCTCGGAGGTTT	31132
Qy	481	CCCTGGGCGGGCAATCGGCGCCATAGCTGCGCTTGCACACCTTCAACCTGAGACATTAACCTT	540
Db	31133	CCCTGGGCTGGCAATAGCTCCTCATAGCTGCGCTGCAATCCCTTAACCTGAGACATTAACCTTC	31192
Qy	541	ATTGACGCACTGCCCCCGAGGCGCTTGTGCCAGCTGATGACTTCTCCGCTGTGACCTTC	600
Db	31193	ATTGATACGCTGTGCCCCGAGGTGTTTTTGGCCAGCACTTACGCACTCTCCCGCTGTGACCTTC	31252
Qy	601	ACCTCCAAACGCGCTGTGCGACGCTGTGCGTCCGGAACCGGCTTTTCTGTGTGGCGGTATGTGA	660
Db	31253	ACCTCCAAACGCGCTGTGCGACCTGTGACCTGTATCTCACTTCTCTGTGGGATGTGGATGTCT	31312
Qy	661	GAGGCGTCTCCGCGCCCCCTGTGTGTGATGACTTTAGCGGGAACCCCTGTGACTGTCAACTGT	720
Db	31313	GAGGCGTCAACCTTCCCTCTGTGTGTGATGACTTTAGCGGGAACCCCACTGTGCAACTGTCT	31372
Qy	721	GAGGTGTGTGTGTGTGTGTGGGGCGCTGTGGCGGGCGGAGACCTGTGAACGTGTGTGTGTCCGG	780
Db	31373	GAGGTGTGTGTGTGTGTGTGGGGCGCTGTGGCGGGCGGAGACCTGTGAACGTGTGTGTGTCCA	31432
Qy	781	CCCGGCGTGTGCGCGGCGGCTACTTCTGTGGCGAGTCCCGGAGGGCGAGTTCTCTGTGTGACCG	840
Db	31433	CCAACCGTGTGAGGCGGCTACTTCTGTGGCGAGTCCGTGAGGGAAGTTCTCTGTGTGACGCT	31492
Qy	841	CCCTCTAATTGCGCGGCACACGCAAGCGCTCTGTGGTGTGTGAAGCCAGGGGCTCAAGCTG	900
Db	31493	CCACTGTATTGCGCGGCACACACAGCGCGCTGTGGTGTGTGAAGGTCAAGGGGCGCACTCTTA	31552
Qy	901	CGGTGTGCGGGCGCTGTGGTGTGACCCCGCGGCTTAACCATGTGACTGTGGTGTGTGTGACACCG	960
Db	31553	CGGTGTGAGGGCTCTGTGGTGTGACCTGTATCTAACAATGTGACTGTGGTGTGTGTGTGACAG	31612
Qy	961	TTGTGTGTGCAACTCTCTCCGAGGCCCGGGCTTTTCCCAACGGGACCTTTAGATTTGGGGTGT	1020
Db	31613	CTGTGTGTGTGCAACTCTTCTCGAGGCTGTGGGCTTTTCCCAACGGGACCTTATGATTTGGGGTGT	31672
Qy	1021	ACCGGCGCTGTGGGCGGCTGTGGGGGCTAACCTGTATGTGCAACCAACCTGTGTGTGTGAGGCT	1080
Db	31673	ACGAGCGCTGTGAGATGTGAGAGGCTTAACCTGTATGTGTCTAACCACTTGTGTGTGTGAGGCT	31732
Qy	1081	ACAGCGCGGATGTGAACTGTGCGGGTGTGTGGCTTGTGCCATGTGTGGGAACAGCACTGTGCGAG	1140

Db	31733	ACAGCCCGAGTAGAGCTCCG369GACTGAGCCTTGCCCACTGATG9GGAACACCA6GTGCTGAG	31792
Qy	1141	GGGGGCGCGCCCGGGGCGCTTGAGACATCGCGCGCTCGGCTCGCACTGCTCGAGGGGTGAG	1200
Db	31793	GGGGGCGCGTCTGGGGCGCTCGACATATGCTGCTTGCTAGAACTGCTCTGAAGGCGAG	31852
Qy	1201	GGAGCGCTGAGACTGAGCCAGCGCGTGCAGAGTGCAGGAGTGCAGCCGCACTTCAGGGCTG	1260
Db	31853	GGAACCTTAGAATCTGAGACCGCGCTGCAGAGTGCAGAGGTGACTGCGCACTCTGAGACTG	31912
Qy	1261	GTGAGCTGGGGGTCCCGGGCGGGCCAGCCGACCGAGTGTGATGTTCCAAATCCAGTACAA	1320
Db	31913	GTGAGCTGGGGGCCCGAGGGCGGCAGCTGACCCAGTGTGATGTTCCAAATCCAGTACAA	31972
Qy	1321	AGCAGCGCAATATGATGACACCTCATCTACCGGATGTGTCCAGCGCTCCAGCCACCACTTCTG	1380
Db	31973	AGCAGCGCAGATGATGACACCTCATCTACCGGTGAGATGGGCGCTTCGACACTCCCTTCAGC	32032
Qy	1381	CTG 1383	
Db	32033	CTG 32035	
RESULT 15			
AC128739			
LOCUS	195342 bp	DNA	linear
DEFINITION	Mus musculus BAC clone RP24-168G14 from chromosome 19, complete sequence.		
ACCESSION	AC128739		
VERSION	AC128739.4	GI:28475716	
KEYWORDS	HTG.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Waligorski, J., Haglund, K. and Bielicki, L.		
TITLE	The sequence of Mus musculus BAC clone RP24-168G14		
JOURNAL	Unpublished (2001)		
REFERENCE	2 (bases 1 to 195342)		
AUTHORS	Wilson, R.		
TITLE	Sequencing of Mus musculus		
JOURNAL	Unpublished (2001)		
REFERENCE	3 (bases 1 to 195342)		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	4 (bases 1 to 195342)		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-FEB-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	5 (bases 1 to 195342)		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	6 (bases 1 to 195342)		
AUTHORS	Wilson, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
COMMENT	On Feb 23, 2003 this sequence version replaced gi:28416270.		
	Genome Center		
	Center: Washington University Genome Sequencing Center		
	Center code: WUGSC		
	Web site: http://genome.wustl.edu		
	Contact: submission@wustl.wustl.edu		
	Summary Statistics		
	Center project name: W_BB0168G14		



NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION: The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION: This sequence is the entire insert of the clone. This clone is overlapped by AC124347.

#### FEATURES

##### source

1. 195342  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="19"  
/map="19"  
/clone="RP24-168G14"  
/clone\_1b="RPCI-24"  
99. . 242  
/rpt\_family="Alu"  
332. . 529  
/rpt\_family="B2"  
576. . 723  
/rpt\_family="Alu"  
583. . 731  
/rpt\_family="B4"  
732. . 870  
/rpt\_family="Alu"  
871. . 945  
/rpt\_family="B4"  
945. . 1065  
/rpt\_family="Alu"  
1073. . 1331  
/rpt\_family="L1"  
1454. . 1800  
/rpt\_family="MaLR"  
2820. . 2962  
/rpt\_family="Alu"  
3003. . 3068  
/rpt\_family="MER53"  
3086. . 3281  
/rpt\_family="B2"  
3557. . 3720  
/rpt\_family="B4"  
3923. . 3982  
/rpt\_family="L1"  
3982. . 4077  
/rpt\_family="B4"  
4226. . 4357  
/rpt\_family="B4"  
5765. . 5912  
/rpt\_family="Alu"  
5937. . 6108  
/rpt\_family="B4"

repeat\_region 6558. . 6905  
/rpt\_family="MaLR"  
repeat\_region 7443. . 7522  
/rpt\_family="L1"  
repeat\_region 7828. . 7976  
/rpt\_family="Alu"  
8124. . 8278  
/rpt\_family="Alu"  
8582. . 8759  
/rpt\_family="B2"  
8895. . 9034  
/rpt\_family="B4"  
9214. . 9317  
/rpt\_family="L1"  
9380. . 9525  
/rpt\_family="Alu"  
9938. . 10114  
/rpt\_family="B4"  
10167. . 10282  
/rpt\_family="Alu"  
10348. . 10479  
/rpt\_family="Alu"  
10543. . 10732  
/rpt\_family="B2"  
11623. . 11746  
/rpt\_family="B4"  
11790. . 11983  
/rpt\_family="L1"  
12002. . 13314  
/rpt\_family="L1"  
13479. . 13559  
/rpt\_family="L1"  
13710. . 14395  
/rpt\_family="RMR19B"  
14403. . 14493  
/rpt\_family="B4"  
15359. . 15506  
/rpt\_family="Alu"  
15712. . 16025  
/rpt\_family="MaLR"  
16270. . 16397  
/rpt\_family="RMR15"  
16501. . 16624  
/rpt\_family="Alu"  
16705. . 16836  
/rpt\_family="ERVK"  
16919. . 16952  
/rpt\_family="ERVK"  
16933. . 17169  
/rpt\_family="ERVK"  
17511. . 17576  
/rpt\_family="Alu"  
17651. . 17779  
/rpt\_family="Alu"  
17780. . 17951  
/rpt\_family="MER1\_type"  
17952. . 17982  
/rpt\_family="B4"  
18633. . 19157  
/rpt\_family="L1"  
19511. . 19639  
/rpt\_family="B4"  
19888. . 20036  
/rpt\_family="L1"  
20081. . 20463  
/rpt\_family="MaLR"  
20467. . 20608  
/rpt\_family="L1"  
20650. . 20826  
/rpt\_family="B2"  
21346. . 21464  
/rpt\_family="B4"  
21531. . 21680



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 17, 2004, 22:32:38 ; Search time 197 Seconds

(without alignments)  
1854.634 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336

Sequence: 1 MAPPLLLLLASGAAACPLP.....LGACRGVGSARLEESVY 635

Scoring table: BLAST62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3336	100.0	635	2	Q6PUG9
2	3336	100.0	635	2	Aah15581
3	3327.5	96.7	636	2	Q80XU8
4	3327.5	96.7	636	2	Q8K3C4
5	3327.5	96.7	636	2	BAC38259
6	2784	83.5	528	2	Aah27475
7	2302.5	69.0	541	2	Q6PK41
8	2302.5	69.0	541	2	Aah07718
9	2013	60.3	460	2	Aah71866
10	1733	60.3	460	2	Aah71866
11	1733	51.9	329	2	Q713C2
12	1733	51.9	329	2	Aah14040
13	1720	51.6	327	2	Q9BWM0
14	1720	51.6	327	2	Aah00207
15	1700	51.0	324	2	Q8N644
16	1662.5	49.8	322	2	Q9UMH4
17	1657.5	49.7	789	2	Q9BE71
18	1644	49.3	788	2	BAC31891
19	1644	49.3	833	2	Q80T99
20	1630	48.9	788	2	Q9CYK3
21	1556	46.6	628	2	Q9BNT0
22	1556	46.6	628	2	BAD18740
23	1546.5	46.4	627	2	QADY10
24	1546.5	46.4	627	2	AAG88494
25	1539.5	46.1	626	2	Q8BLY3
26	1539.5	46.1	626	2	Aah66999
27	1486	44.5	719	1	LRF5_HUMAN
28	1479	44.3	719	1	LRF5_MOUSE
29	1331.5	39.9	700	2	Q9PZ44
30	1293.5	38.8	584	2	Q6PK41
31	1293.5	38.8	584	2	Aah56798

32	1223	36.7	492	2	Q99KT6	Q99KT6 mus musculu
33	1139	34.1	463	2	Q8CIV9	Q8CIV9 mus musculu
34	884	26.5	542	2	Q8TBS9	Q8TBS9 homo sapien
35	618.5	18.5	450	2	Q96C50	Q96C50 homo sapien
36	499	15.0	593	2	Q6UY18	Q6UY18 homo sapien
37	499	15.0	593	2	AAG88482	AAG88482 homo sapi
38	495	14.8	618	2	BAC30759	BAC30759 m adult m
39	482.5	14.5	709	2	Aah60263	Aah60263 mus muscu
40	456.5	13.7	730	2	Q6PMP6	Q6PMP6 mus musculu
41	456.5	13.7	730	2	Aah56458	Aah56458 mus muscu
42	455.5	13.7	730	2	Q6US92	Q6US92 mus musculu
43	455.5	13.7	730	2	AAG74241	AAG74241 mus muscu
44	433.5	13.0	713	2	Q8N182	Q8N182 homo sapien
45	433.5	13.0	713	2	Aah68541	Aah68541 homo sapi

## ALIGNMENTS

RESULT 1

ID Q6PUG9 PRELIMINARY; PRT; 635 AA.

AC Q6PUG9

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Leucine rich repeat and fibronectin type III domain containing 4.

CN Name-LRRN4;

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schier G.D.,

RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Rahey J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RZ [2]

RP SBOUCSE FROM N.A.

RC TISSUE=Lung;

RA Strausberg R.L.

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC015581; Aah15581.2; -

DR InterPro; IPR003961; FN III.

DR InterPro; IPR008957; FN III-like.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003598; IG-like.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR-type.

DR InterPro; IPR003591; LRR-type.

DR Pfam; PF00041; fn3; 1.

DR Pfam; PF00047; Ig; 1.

DR Pfam; PF00560; LRR; 7.

Query Match	Similarity	Score	DB 2	Length	635
Beet	Local	100.0%	100.0%	Pred. No. 6,3e-200	Matches 635
					Conservative 0; Mismatches 0; Indels 0; Gaps 0
DR PRINTS: PR00019; LEURICHRPT.					
DR SMART: SM00060; FN3; 1.					
DR SMART: SM00409; IG; 1.					
DR SMART: SM00408; IGC2; 1.					
DR SMART: SM00082; LRCT; 1.					
DR SMART: SM00369; LRR_Typ; 6.					
DR PROSITE: PS50853; FN3; 1.					
DR PROSITE: PS50835; IG LIKE; 1.					
DR SEQUENCE 635 AA; 66859 MM; 13B6645A17A92EE CRC64;					
Query Match	100.0%	Score 3336;	DB 2;	Length 635;	
Beet Local Similarity	100.0%	Pred. No. 6,3e-200;			
Matches 635;	Conservative	0;	Mismatches	0;	Indels
					Gaps 0;
QY 1	MAPPILLILLASGAACCPVCCQVLSSTLCGHRGLFAPPVNDRTVELRLADNFI	60			
Db 1	MAPPILLILLASGAACCPVCCQVLSSTLCGHRGLFAPPVNDRTVELRLADNFI	60			
QY 61	QALCPDPFRNMTGLVDTLTLSRNAITRIGARAFGDLESIRSLHDGNRLVEIGTSGRGV	120			
Db 61	QALCPDPFRNMTGLVDTLTLSRNAITRIGARAFGDLESIRSLHDGNRLVEIGTSGRGV	120			
QY 121	NIQHLILISGNQIGRIAPGAFDPELESDLDLSYNNLQVPPAGIGAMPALHTLNDNL	180			
Db 121	NIQHLILISGNQIGRIAPGAFDPELESDLDLSYNNLQVPPAGIGAMPALHTLNDNL	180			
QY 181	IDALPPGAFAQLGOLSRJDLTISNRLATLAPDPLFSRGDAEASPAIVLSPGNDLHNC	240			
Db 181	IDALPPGAFAQLGOLSRJDLTISNRLATLAPDPLFSRGDAEASPAIVLSPGNDLHNC	240			
QY 241	ELIMLRRLARDDLETCSPPGLAGRYFWAYVEGFSCPEPLIARHTQRLWLEGORATL	300			
Db 241	ELIMLRRLARDDLETCSPPGLAGRYFWAYVEGFSCPEPLIARHTQRLWLEGORATL	300			
QY 301	RCRLAGDPAPFMHWGPDDRLVGNSSRARAPFNGLEIGVYGAGDAGGYTCIATNPAGEA	360			
Db 301	RCRLAGDPAPFMHWGPDDRLVGNSSRARAPFNGLEIGVYGAGDAGGYTCIATNPAGEA	360			
QY 361	TARVELRLVALPHGNSASSBEGGRPEPSDIAASARTAAEGEGLEBEPVAVQVTEVATSGL	420			
Db 361	TARVELRLVALPHGNSASSBEGGRPEPSDIAASARTAAEGEGLEBEPVAVQVTEVATSGL	420			
QY 421	VSMGGRGADPVMMQIOYNSSEDETLTYRIYPASHHFLKHLVPGADYDCLIALSPA	480			
Db 421	VSMGGRGADPVMMQIOYNSSEDETLTYRIYPASHHFLKHLVPGADYDCLIALSPA	480			
QY 481	AGPSDLTATRLIGCANFSTLPASPPLCHALQAHVLGGTLTVAGVLYVALLVETVALIVR	540			
Db 481	AGPSDLTATRLIGCANFSTLPASPPLCHALQAHVLGGTLTVAGVLYVALLVETVALIVR	540			
QY 541	GRGAGNGGLPLKLSHVQSQTNCGSEPTTKAHPRRPPRPQSSCLDGDACGYARRL	600			
Db 541	GRGAGNGGLPLKLSHVQSQTNCGSEPTTKAHPRRPPRPQSSCLDGDACGYARRL	600			
QY 601	GGAMARRSHSVHGLIGAGCGRGVGSASARLESVV 635				
Db 601	GGAMARRSHSVHGLIGAGCGRGVGSASARLESVV 635				
RESULT 2					
AAH15581	PRELIMINARY;	PRT;	635 AA.		
ID AAH15581;					
AC AAH15581;					
DT 02-MAR-2004 (Tremblrel. 27, Created)					
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)					
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)					
DE Leucine rich repeat and fibronectin type III domain containing 4.					
GN LRPN4.					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
NCBI_TaxID=9606;					
FN [1]					

RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=22386257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G,			
RA	Klausner R.D., Collins F.S., Wagner U., Shemmen C.M., Schuler G.D.,			
RA	Ahtshui S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang Y., Hasel F.,			
RA	Diatchenko L., Marusina K., Farmer A.C., Rubin G.M., Hong I.,			
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheltz T.E.,			
RA	Brownstein M.J., Udell T.B., Toshynski S., Carrincci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Boesk S.A., McEwen P.J., McKernan K.O., Malek U.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,			
RA	Villallon D.K., Mazny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schnutz J., Myers R.M., Butterfield X.S.,			
RA	Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schen J.E.,			
RA	Jones S.J., Mair M.A.;			
RA	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RV	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RA	Strausberg R.;			
RL	Submitted (OCF+2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; BC015581; AAH15581.2; -			
SEQ	SEQUENCE 635 AA; 6685 MW; 13B6645A17A92EE CRC64;			
Query Match	100.0%;	Score 3336;	DB 2;	Length 635;
Best Local Similarity	100.0%;	Pred. No. 8.3e-200;		
Matches 635;	Conservative	0;	Mismatches 0;	Indels 0;
OY	1 MAPPLILLILSLGAAACPCPCVCQNLSSLSTLCARGLLFVPPVNDERTVELRLADNFI 60			
Db	1 MAPPLILLILSLGAAACPCPCVCQNLSSLSTLCARGLLFVPPVNDERTVELRLADNFI 60			
OY	61 QALGPDPFRNTGTVDLTLSRNAITRIGAPAFGDLESRLHLDNQLRYELGTSGIRGV 120			
Db	61 QALGPDPFRNTGTVDLTLSRNAITRIGAPAFGDLESRLHLDNQLRYELGTSGIRGV 120			
OY	121 NLQHLLISGNOLGRIPAGAPDFDLESDLDLDSYNNLRQVPAGIGAMPALHTLNLDHL 180			
Db	121 NLQHLLISGNOLGRIPAGAPDFDLESDLDLDSYNNLRQVPAGIGAMPALHTLNLDHL 180			
OY	181 IDALPFGPAFOIGQLSRDLTTSNRATTLAPDPLFSRGDAEASPAVLVSFGSNPLHCNC 240			
Db	181 IDALPFGPAFOIGQLSRDLTTSNRATTLAPDPLFSRGDAEASPAVLVSFGSNPLHCNC 240			
OY	241 ELIMRLRLARPDDLTCSPPCLARRYWAVVBGEFSCPEPLIAHTOQLMWLEGORATL 300			
Db	241 ELIMRLRLARPDDLTCSPPCLARRYWAVVBGEFSCPEPLIAHTOQLMWLEGORATL 300			
OY	301 RCRAIGDPAFTMHWGWGPPDRIVLVGNSRRARAFAFNGTLEIGVTAGAGAGYTCTIATNPAGEA 360			
Db	301 RCRAIGDPAFTMHWGWGPPDRIVLVGNSRRARAFAFNGTLEIGVTAGAGAGYTCTIATNPAGEA 360			
OY	361 TARVELRYVALPHGNSSSABEGRPCPSDI AASARTPAEGEGTLESEPAVOVTETVATSGI 420			
Db	361 TARVELRYVALPHGNSSSABEGRPCPSDI AASARTPAEGEGTLESEPAVOVTETVATSGI 420			
OY	421 VSMWGGRPADPVMMQIOYNSSEDETLIRIYPASHHFLKGLVPAGDYDCIALASP 480			
Db	421 VSMWGGRPADPVMMQIOYNSSEDETLIRIYPASHHFLKGLVPAGDYDCIALASP 480			
OY	481 AGPSDLTATRLLGCASFSTLPASPLCHALQAATVVGTTLVAVGVVAALLVFVALLVR 540			
Db	481 AGPSDLTATRLLGCASFSTLPASPLCHALQAATVVGTTLVAVGVVAALLVFVALLVR 540			
OY	541 GRGAANGRLPLKLNVQSTNGSBPTKAHPRRSPRRPQSSGLDLDGACCYGARRL 600			
Db	541 GRGAANGRLPLKLNVQSTNGSBPTKAHPRRSPRRPQSSGLDLDGACCYGARRL 600			

Db 541 GRGANGRLPLKLSHVOSQNGBSPPTPKHPPRSPRQSCSLDLDGAGCYGARRL 600  
 QY 601 GGAMARRSHSVHGGILGAGCGVGGSAERLEESVY 635  
 Db 601 GGAMARRSHSVHGGILGAGCGVGGSAERLEESVY 635  
 RESULT 3  
 Q80XU8 PRELIMINARY; PRT; 636 AA.  
 ID 080XU8  
 AC 080XU8;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE leucine rich repeat and fibronectin type III domain containing 4.  
 GN Name=lrtn4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 BX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Pange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC023036; AAH23036.1; -.  
 DR HSSP; Q9BZR6; 1P8T.  
 DR MGD; MGI:2385612; lrtn4.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR008957; FN\_III-like.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR00047; fn3; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00560; LRR; 7.  
 DR PRINTS; PRO0019; LEURICRPT.  
 DR PROSITE; PS00853; FN3; 1.  
 DR PROSITE; PS00853; IG-LIKE; 1.  
 SQ SEQUENCE 636 AA; 67251 MW; 1EE86E96C88BA91 CRC64;

Query Match 96.7%; Score 3227.5; DB 2; Length 636;  
 Best Local Similarity 96.7%; Pred. No. 4,9e-193;  
 Matches 615; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAPPULLLLASGAACPLPCVCQNTSESLSTICAHRLGLFPVPNDVRRVETLADNFI 60  
 Db 1 MAPPULLLLASGAACPLPCVCQNTSESLSTICAHRLGLFPVPNDVRRVETLADNFI 60

QY 61 QALGPPDFRNMGTVDLTLSRNAITRIGARFGLDLSRLHDGNRLVELGTSLRGPV 120  
 Db 61 QALGPPDFRNMGTVDLTLSRNAITRIGARFGLDLSRLHDGNRLVELGTSLRGPV 120  
 QY 121 NLQHLILSGNQLGRIPGAFDDFLSELDLDSYNNLRQYPMWAGIGAMPALHTLNDHNL 180  
 Db 121 NLQHLILSGNQLGRIPGAFDDFLSELDLDSYNNLRQYPMWAGIGAMPALHTLNDHNL 180  
 QY 181 IDALPPGAFAPOLGOSRLDLSNRRLATLAPDPLFSRGRPAEAPALVYSPGNPLHCNC 240  
 Db 181 IDALPPGAFAPOLGOSRLDLSNRRLATLAPDPLFSRGRPAEAPALVYSPGNPLHCNC 240  
 QY 241 ELIMLRRLARPDDLETCAAPGAGRYFMAVEGEGSCPEPLIARTORTLMVLEGORATL 300  
 Db 241 ELIMLRRLARPDDLETCAAPGAGRYFMAVEGEGSCPEPLIARTORTLMVLEGORATL 300  
 QY 301 RCRAIADPPATMHWGPDRLVGNSSRADAPFNGTLEIGVTAGAGAGYTCTATNPAGEA 360  
 Db 301 RCRAIADPPATMHWGPDRLVGNSSRADAPFNGTLEIGVTAGAGAGYTCTATNPAGEA 360  
 QY 361 TARVELRVIALPHGNGSSAEGRPGPSDIAASARTAAEEGTLESFPAVQVTEVATSGI 420  
 Db 361 TARVELRVIALPHGNGSSAEGRPGPSDIAASARTAAEEGTLESFPAVQVTEVATSGI 420  
 QY 421 VSMGPRPADPYWMFQIQNSSDEDTLIYRIVPASSHFLKHLVPGADYDCLIALSPA 480  
 Db 421 VSMGPRPADPYWMFQIQNSSDEDTLIYRIVPASSHFLKHLVPGADYDCLIALSPA 480  
 QY 481 AGPSDITATRLIGCAHFSTLPASPLCHALQAHVGGTLTVAVGVIVAAIVFTVALVR 540  
 Db 481 AGPSDITATRLIGCAHFSTLPASPLCHALQAHVGGTLTVAVGVIVAAIVFTVALVR 540  
 QY 541 GKGAGNGRLPLKLSHVOSQNGSPPTPKHPPRSPRQSCSLDLDGAGCYGAR 599  
 Db 541 GKGAGNGRLPLKLSHVOSQNGSPPTPKHPPRSPRQSCSLDLDGAGCYGAR 600  
 QY 600 LGAMARRSHSVHGGILGAGCGVGGSAERLEESVY 635  
 Db 600 LGAMARRSHSVHGGILGAGCGVGGSAERLEESVY 635

RESULT 4  
 Q8K3C4 PRELIMINARY; PRT; 636 AA.  
 ID 08K3C4  
 AC 08K3C4;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE lrtn4 protein.  
 GN Name=lrtn4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 BX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Pange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzyzinski M.I., Skaleka U., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RA Strausberg R.; (Feb-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; BC023156; AAR23156.1; -  
 DR HSBP; Q9BZR6; 1P8T.  
 DR MGD; MGI:2385612; Lrfn4.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_Typ.  
 DR Pfam; PF00041; fn3; 1.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00560; LRR; 7.  
 DR PRINTS; PR00019; LEUR1CHRP1.  
 DR SMART; SM00060; FN3; 1.  
 DR SMART; SM00408; IGG2; 1.  
 DR SMART; SM00082; LRRT; 1.  
 DR SMART; SM00013; LRRT; 1.  
 DR SMART; SM00369; LRR\_Typ; 2.  
 DR SMART; PS50853; FN3; 1.  
 DR PROSITE; PS50835; IG\_Like; 1.  
 SQ SEQUENCE 636 AA; 67264 MW; 9CE86E8A3981A884 CRC64;  
 Query Match 96.7%; Score 3227.5; DB 2; Length 636;  
 Best Local Similarity 96.7%; Pred. No. 4.9e-193;  
 Matches 615; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAPPILLILLASGAACPLPCVCONLSSTLCANHGGLFVPPNVRRFVEYRIADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCONLSSTLCANHGGLFVPPNVRRFVEYRIADNFI 60  
 QY 61 QALGPPDFRMTGLVDLTLSRNATIRIGARFGDLESLRSLHDGNRLVELGTSLRGPV 120  
 DB 61 QALGPPDFRMTGLVDLTLSRNATIRIGARFGDLESLRSLHDGNRLVELGTSLRGPV 120  
 QY 61 QALGPPDFRMTGLVDLTLSRNATIRIGARFGDLESLRSLHDGNRLVELGTSLRGPV 120  
 DB 61 QALGPPDFRMTGLVDLTLSRNATIRIGARFGDLESLRSLHDGNRLVELGTSLRGPV 120  
 QY 121 NLGHLITSSGQIGRIAGADPDFLESIEDLSTNNLRQVPMAGIGAMPALHTINDHNL 180  
 DB 121 NLGHLITSSGQIGRIAGADPDFLESIEDLSTNNLRQVPMAGIGAMPALHTINDHNL 180  
 QY 121 NLGHLITSSGQIGRIAGADPDFLESIEDLSTNNLRQVPMAGIGAMPALHTINDHNL 180  
 DB 121 NLGHLITSSGQIGRIAGADPDFLESIEDLSTNNLRQVPMAGIGAMPALHTINDHNL 180  
 QY 181 IDALPPGAFQOLGSLDLTSLNRLATLAPDPLPSRGRDAEAPAPVLSFSGNPLHCNC 240  
 DB 181 IDALPPGAFQOLGSLDLTSLNRLATLAPDPLPSRGRDAEAPAPVLSFSGNPLHCNC 240  
 QY 181 IDALPPGAFQOLGSLDLTSLNRLATLAPDPLPSRGRDAEAPAPVLSFSGNPLHCNC 240  
 DB 181 IDALPPGAFQOLGSLDLTSLNRLATLAPDPLPSRGRDAEAPAPVLSFSGNPLHCNC 240  
 QY 241 ELIMLRRLAPDDLETGASPPGLAGRYFWAVPEGEFSCEPPLIARHTQRLMWLEGQRATL 300  
 DB 241 ELIMLRRLAPDDLETGASPPGLAGRYFWAVPEGEFSCEPPLIARHTQRLMWLEGQRATL 300  
 QY 241 ELIMLRRLAPDDLETGASPPGLAGRYFWAVPEGEFSCEPPLIARHTQRLMWLEGQRATL 300  
 DB 241 ELIMLRRLAPDDLETGASPPGLAGRYFWAVPEGEFSCEPPLIARHTQRLMWLEGQRATL 300  
 QY 301 RCRAISGPAAPPMHVGDDDLVGNSSRAAPFNPTLEIGYTGADAGGYCIAINPGEA 360  
 DB 301 RCRAISGPAAPPMHVGDDDLVGNSSRAAPFNPTLEIGYTGADAGGYCIAINPGEA 360  
 QY 301 RCRAISGPAAPPMHVGDDDLVGNSSRAAPFNPTLEIGYTGADAGGYCIAINPGEA 360  
 DB 301 RCRAISGPAAPPMHVGDDDLVGNSSRAAPFNPTLEIGYTGADAGGYCIAINPGEA 360  
 QY 361 TARVELLVLAIPHGNGSSAEGRPGSPDIASARTAAEGEGTLESBPVQTEVTATSGL 420  
 DB 361 TARVELLVLAIPHGNGSSAEGRPGSPDIASARTAAEGEGTLESBPVQTEVTATSGL 420  
 QY 361 TARVELLVLAIPHGNGSSAEGRPGSPDIASARTAAEGEGTLESBPVQTEVTATSGL 420  
 DB 361 TARVELLVLAIPHGNGSSAEGRPGSPDIASARTAAEGEGTLESBPVQTEVTATSGL 420  
 QY 421 VSWGPGPADPVMVFOIYNNSEDETLIYRIVPASHHFLKHLVPADYDLCTLAISPA 480  
 DB 421 VSWGPGPADPVMVFOIYNNSEDETLIYRIVPASHHFLKHLVPADYDLCTLAISPA 480  
 QY 421 VSWGPGPADPVMVFOIYNNSEDETLIYRIVPASHHFLKHLVPADYDLCTLAISPA 480  
 DB 421 VSWGPGPADPVMVFOIYNNSEDETLIYRIVPASHHFLKHLVPADYDLCTLAISPA 480  
 QY 481 AGRSDDLTAATLLGCAHSTIPASPLCHALQAHVIGTLTAAVGVVYAAILVFVALLVR 540  
 DB 481 AGRSDDLTAATLLGCAHSTIPASPLCHALQAHVIGTLTAAVGVVYAAILVFVALLVR 540  
 QY 481 AGRSDDLTAATLLGCAHSTIPASPLCHALQAHVIGTLTAAVGVVYAAILVFVALLVR 540  
 DB 481 AGRSDDLTAATLLGCAHSTIPASPLCHALQAHVIGTLTAAVGVVYAAILVFVALLVR 540

QY 541 GRGANGRLPLKLSHVOSQNGGSPPTKAPHPRSPRPSRSCSIDLGD-AGCYGARR 599  
 DB 541 GRGANGRLPLKLSHVOSQNGGSPPTKAPHPRSPRPSRSCSIDLGD-AGCYGARR 600  
 QY 600 LGGAMARRSHSVHGLLGAGCRGVGSGARLESVV 635  
 DB 601 LGGAMARRSHSVHGLLGAGCRGVGSGARLESVV 636  
 RESULT 5  
 BAC38259  
 ID BAC38259 PRELIMINARY; PRT; 636 AA.  
 AC BAC38259;  
 DT 14-APR-2004 (TEMBLrel. 27, Created)  
 DT 14-APR-2004 (TEMBLrel. 27, Last sequence update)  
 DT 14-APR-2004 (TEMBLrel. 27, Last annotation update)  
 DE 16 days embryo head cDNA, RIKEN full-length enriched library,  
 DE clone: C30040J16 product: hypothetical Immunoglobulin and major  
 DE histocompatibility complex domain/leucine-rich repeat/fibronectin type  
 DE III domain/ Cysteine-rich flanking region, C-terminal containing  
 DE protein, full insert sequence.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=22354683; PubMed=12466951;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=2049374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Konno H., Akiyama J., Nishii K., Kitunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa S., Ohara E., Wataniki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771 (2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.



RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiroaka T., Hirose T.,  
 RA Hori F., Imocani K., Iehi Y., Ito M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Komoto H., Kouda K., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numata K., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Shibata K., Shinagawa A., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Soabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toyota T., Yasunishi A., Yamamoto S., Hayashizaki Y.,  
 RA Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 KW EMBL; AK081560; BAC8259.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 636 AA; 67251 MW; 1EE86E96CB88BA91 CRC64;  
 Query Match 96.7%; Score 3227.5; DB 2; Length 636;  
 Best Local Similarity 96.7%; Pred. No. 4,9e-193;  
 Matches 615; Conservative 9; Mismatches 11; Indels 1; Gaps 1;  
 QY 1 MAPPRLILLASGAACPLPCVCNLSSTLCARHGLFVPPVDRRTVELRLADNFI 60  
 DB 1 MAPPRLILLASGAACPLPCVCNLSSTLCARHGLFVPPVDRRTVELRLADNFI 60  
 QY 61 QALGPPDFRMTGLVLTLSRNAITRIGAPAFGDLSELRSLHDGNRLVELGTSLRGPV 120  
 DB 61 QALGPPDFRMTGLVLTLSRNAITRIGAPAFGDLSELRSLHDGNRLVELGTSLRGPV 120  
 QY 121 NIQHLISGNQGRIRAPAFDDLESLDLSYNNLRQVPAGIGAMPRLHNLNDHNI 180  
 DB 121 NIQHLISGNQGRIRAPAFDDLESLDLSYNNLRQVPAGIGAMPRLHNLNDHNI 180  
 QY 181 IDALPPGAFQOLGQLSRDLTNSRLATLAPDPLFSRGRDAEAPAPLVISFGNPLHCNC 240  
 DB 181 IDALPPGAFQOLGQLSRDLTNSRLATLAPDPLFSRGRDAEAPAPLVISFGNPLHCNC 240  
 QY 241 ELLMRLRLRPDDLETCASPPGLAGRYFMAVPEGEFSCBEPPIARHTORLWLEGORATL 300  
 DB 241 ELLMRLRLRPDDLETCASPPGLAGRYFMAVPEGEFSCBEPPIARHTORLWLEGORATL 300  
 QY 301 RCRALGDPAPMTMWVPPDRLVGNSSRAAFAPFNGTLEIGVTGAGDAGYTCIATNPAGA 360  
 DB 301 RCRALGDPAPMTMWVPPDRLVGNSSRAAFAPFNGTLEIGVTGAGDAGYTCIATNPAGA 360  
 QY 361 TRAVELRVIALPHGNSAEGRPSPSDIAASARTABEGGTLESPPAVQVTEVATSGL 420  
 DB 361 TRAVELRVIALPHGNSAEGRPSPSDIAASARTABEGGTLESPPAVQVTEVATSGL 420  
 QY 421 VSMGGRPADPVMWFQIYNSSSEDETLIYRIYVASSHFLKHLVPGADYDCLIALSPA 480  
 DB 421 VSMGGRPADPVMWFQIYNSSSEDETLIYRIYVASSHFLKHLVPGADYDCLIALSPA 480  
 QY 481 AGPSDLTARLLGCAHFTLPSPLCHALQAVLGGTLTVAVGVVVAALLVFTVALVR 540  
 DB 481 AGPSDLTARLLGCAHFTLPSPLCHALQAVLGGTLTVAVGVVVAALLVFTVALVR 540  
 QY 541 GRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCLDLID AGCYGYARR 599  
 DB 541 GRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCLDLID AGCYGYARR 599  
 QY 600 LGGAMARRSHSVHGLILGAGCGVGSABRLTESYV 635  
 DB 600 LGGAMARRSHSVHGLILGAGCGVGSABRLTESYV 635  
 QY 601 LGGAMARRSHSVHGLILGAGCGVGSABRLTESYV 636  
 DB 601 LGGAMARRSHSVHGLILGAGCGVGSABRLTESYV 636

DE LRFN4 protein (Fragment).  
 GN LRFN4.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung.  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner F.S., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,  
 RA Diatchenko L., Matsushita K., Farnier A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shychenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schermer A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung.  
 RA Strausberg R.,  
 RA Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC027475; AAH27475.2; -  
 FT NON TER 1  
 SQ SEQUENCE 528 AA; 55315 MW; E5B57A25FF22F2B2 CRC64;  
 Query Match 83.5%; Score 2784; DB 2; Length 528;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-165;  
 Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 108 LVELGTSLRGVNIQHLISGNQGRIRAPAFDDLESLDLSYNNLRQVPAGIGTA 167  
 DB 1 LVELGTSLRGVNIQHLISGNQGRIRAPAFDDLESLDLSYNNLRQVPAGIGTA 167  
 QY 168 MPALHTLNDHNLIDALPPGAFQOLGQLSRDLTNSRLATLAPDPLFSRGRDAEAPAPLV 227  
 DB 168 MPALHTLNDHNLIDALPPGAFQOLGQLSRDLTNSRLATLAPDPLFSRGRDAEAPAPLV 227  
 QY 228 VLSFSGNPLHCNCCELLMLRLRPDDLETCASPPGLAGRYFMAVPEGEFSCBEPPIARHT 287  
 DB 228 VLSFSGNPLHCNCCELLMLRLRPDDLETCASPPGLAGRYFMAVPEGEFSCBEPPIARHT 287  
 QY 288 GRMTVLBGRATLRCRALGDPAPMTMWVPPDRLVGNSSRAAFAPFNGTLEIGVTGAGDAG 347  
 DB 288 GRMTVLBGRATLRCRALGDPAPMTMWVPPDRLVGNSSRAAFAPFNGTLEIGVTGAGDAG 347  
 QY 348 GYTCTIATNPAGEATARVELRVIALPHGNSAEGRPSPSDIAASARTABEGGTLESPP 407  
 DB 348 GYTCTIATNPAGEATARVELRVIALPHGNSAEGRPSPSDIAASARTABEGGTLESPP 407  
 QY 408 AVQVTEVATSGLVSWGGRPADPVMWFQIYNSSSEDETLIYRIYVASSHFLKHLVPG 467  
 DB 408 AVQVTEVATSGLVSWGGRPADPVMWFQIYNSSSEDETLIYRIYVASSHFLKHLVPG 467  
 QY 468 ADVDCIALSPAPSPDLTARLLGCAHFTLPSPLCHALQAVLGGTLTVAVGVVVAALLVFTVALVR 527  
 DB 468 ADVDCIALSPAPSPDLTARLLGCAHFTLPSPLCHALQAVLGGTLTVAVGVVVAALLVFTVALVR 527  
 QY 528 AALLVFTVALVRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCLDLID 587  
 DB 528 AALLVFTVALVRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCLDLID 587

Db 421 AALIVTVALVIRGAGNGRLPLKLSHVOSQTNGBSPTRKXAPPRPPRQSCSLD 480

QY 588 LGDAGCYVARRLGAMARRSHSVHGGLLGAGCRGVGGSARLEESV 635

Db 481 LGDAGCYVARRLGAMARRSHSVHGGLLGAGCRGVGGSARLEESV 528

RESULT 7

Q6PK41 ID Q6PK41 PRELIMINARY; PRT; 541 AA.

AC Q6PK41; 05-JUL-2004 (TEMBLrel. 27, Created)

DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)

DE LRFN4 protein (fragment).

GN Name=LRFN4;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Muscle;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RL [2]

RN SEQUENCE FROM N.A.

RP TISSUE=Muscle;

RC Strausberg R.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC007718; AAH07718.1; -like.

DR InterPro; IPR008957; FW\_III-like.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003598; IG\_c2.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_cterm.

DR InterPro; IPR003591; LRR\_cyp.

DR Pfam; PF00047; IG\_1.

DR Pfam; PF00560; LRR\_7.

DR PRINTS; PR0019; LEURICHRPT.

DR SMART; SM00409; IG\_1.

DR SMART; SM00408; IG\_c2; 1.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_Typ; 6.

DR PROSITE; PS50835; IG\_LIKE; 1.

FT NON TER 541 541

SQ SEQUENCE 541 AA; 58167 MW; E01B98549DFE8E4F CRC64;

Query March 69.0%; Score 2302.5; DB 2; Length 541;

Best-Local Similarity 90.9%; Pred. No. 2e-135; Indels 19; Gaps 3;

Matches 450; Conservative 5; Mismatches 21;

1 MAPPLILLILASGAACPICVCQNLSESLTLCARHGLTFPPNVDRRTVELRLADNFI 60

Db 1 MAPPLILLILASGAACPICVCQNLSESLTLCARHGLTFPPNVDRRTVELRLADNFI 60

QY 61 QALGPPDFNNMTGLVDLTLSRNAITTRIGARAFDLESLSLHIDNKRVLVLTGSLRGV 120

Db 61 QALGPPDFNNMTGLVDLTLSRNAITTRIGARAFDLESLSLHIDNKRVLVLTGSLRGV 120

QY 121 NIQHLTILSGNQLGRIPGAFDDPLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHNL 180

Db 121 NIQHLTILSGNQLGRIPGAFDDPLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHNL 180

QY 181 IDALPPGAPAOIGQLSRLDLTNSRLATLAPDPLFSRGRDAEAPLVISFGNPLHCNC 240

Db 181 IDALPPGAPAOIGQLSRLDLTNSRLATLAPDPLFSRGRDAEAPLVISFGNPLHCNC 240

QY 241 ELLIMRLARPDDLEFCASPPGLAGYFMAVPEGESCEPPLIARHTQRLMTVEGORATL 300

Db 241 ELLIMRLARPDDLEFCASPPGLAGYFMAVPEGESCEPPLIARHTQRLMTVEGORATL 300

QY 301 RCRAIDGPAPVPMVGPDDDLVGNSSRARAFPMGTLEIGVTGAGDAGYTCIATNPAGEA 360

Db 301 RCRAIDGPAPVPMVGPDDDLVGNSSRARAFPMGTLEIGVTGAGDAGYTCIATNPAGEA 360

QY 361 TAEVELRVIALPHGNSSAEGARPPSPDIAASARTAAEGEGLSEEPVAVTEVTATGCL 420

Db 361 TAEVELRVIALPHGNSSAEGARPPSPDIAASARTAAEGEGLSEEPVAVTEVTATGCL 420

QY 421 VSMGPRPADPVMVPMQIYNNSEDETLTYIVPASHHFLKTLVGCADVDLCLLALSP- 479

Db 421 VSMGPRPADPVMVPMQIYNNSEDETLTYIVPASHHFLKTLVGCADVDLCLLALSP- 479

QY 480 -----AAGPSD 485

Db 472 LGSGLQTRVAPGRSN 486

RESULT 8

AAH07718 ID AAH07718 PRELIMINARY; PRT; 541 AA.

AC AAH07718;

DT 02-MAR-2004 (TEMBLrel. 27, Created)

DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)

DE LRFN4 protein (fragment).

GN LRFN4

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Muscle;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RL Strausberg R.;  
 DR Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL: BC007718; AA07718.1; -  
 FT NON\_TER  
 SQ SEQUENCE 541 AA; 58167 MW; E01B98549DF6BE4F CRC64;  
  
 Query Match 69.0%; Score 2302.5; DB 2; Length 541;  
 Best Local Similarity 90.9%; Pred. No. 2e-135;  
 Matches 450; Conservative 5; Mismatches 21; Indels 19; Gaps 3;  
  
 QY 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 QY 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHLDGNRLVELGTGSLRGPV 120  
 DB 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHLDGNRLVELGTGSLRGPV 120  
 QY 121 NIQHLITSGNQLGRIAPGAFDDFLESLEDDLSYNNLRQVPWAGIGAMPALHTLNDHNL 180  
 DB 121 NIQHLITSGNQLGRIAPGAFDDFLESLEDDLSYNNLRQVPWAGIGAMPALHTLNDHNL 180  
 QY 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 DB 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 QY 241 ELLMRLRLARPDLETCASPPGLAGRYFMAVBEGESCEPPLIARTQRLMWLEGQRATL 300  
 DB 241 ELLMRLRLARPDLETCASPPGLAGRYFMAVBEGESCEPPLIARTQRLMWLEGQRATL 300  
 QY 301 RCRAIDPAPTMHWGPDDRLVGNSSRAAFPNGTLEIGVTAGAGDGATCTATNPAGEA 360  
 DB 301 RCRAIDPAPTMHWGPDDRLVGNSSRAAFPNGTLEIGVTAGAGDGATCTATNPAGEA 360  
 QY 361 TARVELRVIALPHGNSAEGGRPGSPDIAASARTAEGBGTSESPAVOTEVTAISGL 420  
 DB 361 TARVELRVIALPHGNSAEGGRPGSPDIAASARTAEGBGTSESPAVOTEVTAISGL 420  
 QY 421 VSKGPRPADPVMFQIOTNNSSEDEFTLIRIVPASHHPLKHLVQADYDLCLALSP 479  
 DB 421 VSKGPRPADPVMFQIOTNNSSEDEFTLIRIVPASHHPLKHLVQADYDLCLALSP 479  
 QY 480 -----AAGPSD 485  
 DB 472 LSGGLQTRVYAGPSN 486  
  
 RESULT 9  
 Q6IPL6 PRELIMINARY; PRT; 460 AA.  
 AC 06IPL6;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 OS Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI Taxid=9606;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Spennemann C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin C.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Rane J., Helton B., Kettelman W., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Gittwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Maira M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=skin;  
 RL Strausberg R.;  
 DR Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC071866; AA071866.1; -  
 DR InterPro: IPR003599; IG.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003598; IG\_c2.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Cterm.  
 DR InterPro: IPR003591; LRR\_Typ.  
 DR Pfam: PF00047; IG\_1.  
 DR Pfam: PF00560; LRR\_7.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00409; IG\_1.  
 DR SMART: SM00408; IG\_c2; 1.  
 DR SMART: SM00082; LRRCT; 1.  
 DR SMART: SM00369; LRR\_Typ; 6.  
 DR PROSITE: PS00835; IG\_Like; 1.  
 DR Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 460 AA; 49273 MW; FBD686F0AD360040 CRC64;  
  
 Query Match 60.3%; Score 2013; DB 2; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-117;  
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 QY 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHLDGNRLVELGTGSLRGPV 120  
 DB 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHLDGNRLVELGTGSLRGPV 120  
 QY 121 NIQHLITSGNQLGRIAPGAFDDFLESLEDDLSYNNLRQVPWAGIGAMPALHTLNDHNL 180  
 DB 121 NIQHLITSGNQLGRIAPGAFDDFLESLEDDLSYNNLRQVPWAGIGAMPALHTLNDHNL 180  
 QY 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 DB 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 QY 241 ELLMRLRLARPDLETCASPPGLAGRYFMAVBEGESCEPPLIARTQRLMWLEGQRATL 300  
 DB 241 ELLMRLRLARPDLETCASPPGLAGRYFMAVBEGESCEPPLIARTQRLMWLEGQRATL 300  
 QY 301 RCRAIDPAPTMHWGPDDRLVGNSSRAAFPNGTLEIGVTAGAGDGATCTATNPAGEA 360  
 DB 301 RCRAIDPAPTMHWGPDDRLVGNSSRAAFPNGTLEIGVTAGAGDGATCTATNPAGEA 360  
 QY 361 TARVELRVIALPHGNSAEGGR 383  
 DB 361 TARVELRVIALPHGNSAEGGR 383  
  
 RESULT 10  
 AA071866 PRELIMINARY; PRT; 460 AA.  
 ID AA071866

```

AC AAH71866;
DT 01-JUN-2004 (TReMBLrel. 27, Created)
DT 01-JUN-2004 (TReMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant P., Prange C.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywnicki M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071866; AAH71866.1;
KW Hypothetical protein.
FT NON TER
SQ
SEQUENCE 460 AA; 49273 MW; FBD686F0AD360040 CRC64;

Query Match 60.3%; Score 2013; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 1,9e-117;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPPLILLIAGAAACPLPCVCONLSESLSTCAHGLLFPVNDRTVELRLADNFI 60
DB 1 MAPPLILLIAGAAACPLPCVCONLSESLSTCAHGLLFPVNDRTVELRLADNFI 60
QY 61 QALGPPFRMTGTVDLTLSRNAITRGARAFGLSRLSHLDGNLVLTGCSLRGPV 120
DB 61 QALGPPFRMTGTVDLTLSRNAITRGARAFGLSRLSHLDGNLVLTGCSLRGPV 120
QY 61 QALGPPFRMTGTVDLTLSRNAITRGARAFGLSRLSHLDGNLVLTGCSLRGPV 120
DB 61 QALGPPFRMTGTVDLTLSRNAITRGARAFGLSRLSHLDGNLVLTGCSLRGPV 120
QY 121 NLGHLISGNQIGRIAGAFDDFLSLEDLSTYNNLRQVPMAGMPALHTLNDHNL 180
DB 121 NLGHLISGNQIGRIAGAFDDFLSLEDLSTYNNLRQVPMAGMPALHTLNDHNL 180
QY 121 NLGHLISGNQIGRIAGAFDDFLSLEDLSTYNNLRQVPMAGMPALHTLNDHNL 180
DB 121 NLGHLISGNQIGRIAGAFDDFLSLEDLSTYNNLRQVPMAGMPALHTLNDHNL 180
QY 181 IDALPFAFQOLGSLDLSNRATLAPDPLFSRGRDABAPAPLVISFGNPLHCNC 240
DB 181 IDALPFAFQOLGSLDLSNRATLAPDPLFSRGRDABAPAPLVISFGNPLHCNC 240
QY 181 IDALPFAFQOLGSLDLSNRATLAPDPLFSRGRDABAPAPLVISFGNPLHCNC 240
DB 181 IDALPFAFQOLGSLDLSNRATLAPDPLFSRGRDABAPAPLVISFGNPLHCNC 240
QY 241 ELMLRLRLAPDDLETASPPGLAGRFMAVPSSEFCEPPLTARHQRMLVLEGQATL 300
DB 241 ELMLRLRLAPDDLETASPPGLAGRFMAVPSSEFCEPPLTARHQRMLVLEGQATL 300
QY 241 ELMLRLRLAPDDLETASPPGLAGRFMAVPSSEFCEPPLTARHQRMLVLEGQATL 300
DB 241 ELMLRLRLAPDDLETASPPGLAGRFMAVPSSEFCEPPLTARHQRMLVLEGQATL 300
QY 301 RCALGDPAPETMVGDDRLVGNSSRARAFAFNGTLEIGVTGADAGYTCIATNPAGEA 360
DB 301 RCALGDPAPETMVGDDRLVGNSSRARAFAFNGTLEIGVTGADAGYTCIATNPAGEA 360
QY 301 RCALGDPAPETMVGDDRLVGNSSRARAFAFNGTLEIGVTGADAGYTCIATNPAGEA 360
DB 301 RCALGDPAPETMVGDDRLVGNSSRARAFAFNGTLEIGVTGADAGYTCIATNPAGEA 360
QY 361 TARVELRVTLALPHGNSAEGGR 383
DB 361 TARVELRVTLALPHGNSAEGGR 383

```

```

DB 361 TARVELRVTLALPHGNSAEGGR 383

RESULT 11
Q713C2 PRELIMINARY; PRT; 329 AA.
AC Q713C2;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE LRPN4 protein (Fragment).
OS Homo sapiens (Human).
GN Name=LRPN4;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant P., Prange C.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywnicki M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014040; AAH14040.2;
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PSS0853; FN3; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON TER
SQ
SEQUENCE 329 AA; 33674 MW; 0E78B9038E67DB8 CRC64;

Query Match 51.9%; Score 1733; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.8e-100;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 DPAFTMHWGPPDRLVGNSSRARAFAFNGTLEIGVTGADAGYTCIATNPAGEATRVEL 366
DB 1 DPAFTMHWGPPDRLVGNSSRARAFAFNGTLEIGVTGADAGYTCIATNPAGEATRVEL 366
QY 367 RVIALPHGNSAEGGRPGSDIAAGARTABEGTLESPAVQVBTATSGLVSKPG 426
DB 61 RVIALPHGNSAEGGRPGSDIAAGARTABEGTLESPAVQVBTATSGLVSKPG 426
QY 427 RPADPYWMPQIOYNSSEDFLIRYIVPASGHHFLKHLVGAUDYDCLALSPAAGPSDL 486
DB 121 RPADPYWMPQIOYNSSEDFLIRYIVPASGHHFLKHLVGAUDYDCLALSPAAGPSDL 486
QY 487 TARLLGCAHFTLTPASPPLCHALQAHVLGTLTVAVGVLVVALLVTVALLVGRGAGN 546
DB 487 TARLLGCAHFTLTPASPPLCHALQAHVLGTLTVAVGVLVVALLVTVALLVGRGAGN 546

```

```

Db      181 TATRLGCAHFSTLPASPLCHALQAHVLTVAAGVVAALVFTVALLVRGAGN 240
Qy      547 GRPLKLSHVOSQTNQSPPTPKAHPSPRPORSCSLDLDAGCGYARBLGAMAR 606
Db      241 GRPLKLSHVOSQTNQSPPTPKAHPSPRPORSCSLDLDAGCGYARBLGAMAR 300
Qy      607 RSHSVHGILLGAGCGVGGSAERLEESV 635
Db      301 RSHSVHGILLGAGCGVGGSAERLEESV 329

RESULT 12
AAH14040
ID AAH14040 PRELIMINARY; PRT; 329 AA.
AC AAH14040;
DT 02-MAR-2004 (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE LRFN4 protein (Fragment).
GN LRFN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014040; AAH14040.2; -.
FT NON TER
SQ SEQUENCE 329 AA; 33674 MW; 0E78B9038BE67DB8 CRC64;

Query Match 51.9%; Score 1733; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.8e-100; Indels 0; Gaps 0;
Matches 329; Conservative 0; Mismatches 0;

Db      307 DPAPTMHWGPDRLVGNSSRRARAFPNGLTEIGVTGAGDAGGYTCIATNPAGEATARVEL 366
Qy      367 RVIALPHGNSASBEGRPQSPDIASARTAAEGTLESEPVQVTEVTATSGLVSWGPG 426
Db      61 RVIALPHGNSASBEGRPQSPDIASARTAAEGTLESEPVQVTEVTATSGLVSWGPG 120
Qy      427 RPADPVMVFOIQYNSSEDETLIRIVPASSHHFLKHLVPGADYDLCLLASPAAGPSDL 486
Db      121 RPADPVMVFOIQYNSSEDETLIRIVPASSHHFLKHLVPGADYDLCLLASPAAGPSDL 180

```

```

Qy      487 TATRLGCAHFSTLPASPLCHALQAHVLTVAAGVVAALVFTVALLVRGAGN 546
Db      181 TATRLGCAHFSTLPASPLCHALQAHVLTVAAGVVAALVFTVALLVRGAGN 240
Qy      547 GRPLKLSHVOSQTNQSPPTPKAHPSPRPORSCSLDLDAGCGYARBLGAMAR 606
Db      241 GRPLKLSHVOSQTNQSPPTPKAHPSPRPORSCSLDLDAGCGYARBLGAMAR 300
Qy      607 RSHSVHGILLGAGCGVGGSAERLEESV 635
Db      301 RSHSVHGILLGAGCGVGGSAERLEESV 329

RESULT 13
Q9BMU0
ID Q9BMU0 PRELIMINARY; PRT; 327 AA.
AC Q9BMU0;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE LRFN4 protein (Fragment).
GN LRFN4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC000207; AAH00207.2; -.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR008957; FN_III-like.
DR InterPro: IPR007110; IG-like.
DR Pfam: PF00041; fn3; 1.
DR SMART: SMO0060; FN3; 1.
DR PROSITE: PS0053; FN3; 1.
FT NON TER
SQ SEQUENCE 327 AA; 33461 MW; A5582938BF197214 CRC64;

Query Match 51.6%; Score 1720; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 2.4e-99; Indels 0; Gaps 0;
Matches 327; Conservative 0; Mismatches 0;

Db      309 APPTMHWGPDRLVGNSSRRARAFPNGLTEIGVTGAGDAGGYTCIATNPAGEATARVEL 368
Qy      1 APPTMHWGPDRLVGNSSRRARAFPNGLTEIGVTGAGDAGGYTCIATNPAGEATARVEL 60

```

```

QY 369 LALPHGSSSABGRPGSPDIAASARTAAEGGTLESBAVQVTEYTATSGVSWGPGRP 428
    |||
DB 61 LALPHGSSSABGRPGSPDIAASARTAAEGGTLESBAVQVTEYTATSGVSWGPGRP 120
    |||
QY 429 ADPVMMFOIQVNSSEDETLIYRIVPASSHFLKHLVPGADYDLCILALSPAAGSDLTAA 488
    |||
DB 121 ADPVMMFOIQVNSSEDETLIYRIVPASSHFLKHLVPGADYDLCILALSPAAGSDLTAA 180
    |||
QY 489 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 548
    |||
DB 181 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 240
    |||
QY 549 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 608
    |||
DB 241 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 300
    |||
QY 609 HSYHGGLGAGCRGVGSAERLEESVY 635
    |||
DB 301 HSYHGGLGAGCRGVGSAERLEESVY 327
    |||

```

## RESULT 14

```

AAH00207 PRELIMINARY; PRT; 327 AA.
ID AAH00207
AC AAH00207

```

```

DT 02-MAR-2004 (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE LRFN4 protein (Fragment).
GN LRFN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Manisina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywnski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000207; AAH00207.2; -.
FT NON TER 1
SQ SEQUENCE 327 AA; 33461 MW; A5582938BF197214 CRC64;

```

```

Query Match 51.6%; Score 1720; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. NO. 2.4e-99;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
309. APTMHWGPDRLVGNSSRARAFNGTLEIGVTGAGDAGGYTCIATNPAGATARVELRV 368

```

```

DB 1 APTMHWGPDRLVGNSSRARAFNGTLEIGVTGAGDAGGYTCIATNPAGATARVELRV 60
    |||
QY 369 LALPHGSSSABGRPGSPDIAASARTAAEGGTLESBAVQVTEYTATSGVSWGPGRP 428
    |||
DB 61 LALPHGSSSABGRPGSPDIAASARTAAEGGTLESBAVQVTEYTATSGVSWGPGRP 120
    |||
QY 429 ADPVMMFOIQVNSSEDETLIYRIVPASSHFLKHLVPGADYDLCILALSPAAGSDLTAA 488
    |||
DB 121 ADPVMMFOIQVNSSEDETLIYRIVPASSHFLKHLVPGADYDLCILALSPAAGSDLTAA 180
    |||
QY 489 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 548
    |||
DB 181 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 240
    |||
QY 549 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 608
    |||
DB 241 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 300
    |||
QY 609 HSYHGGLGAGCRGVGSAERLEESVY 635
    |||
DB 301 HSYHGGLGAGCRGVGSAERLEESVY 327
    |||

```

```

RESULT 15
ID 08N644 PRELIMINARY; PRT; 324 AA.
AC 08N644

```

```

DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE LRFN4 protein.
GN Name=LRFN4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Manisina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywnski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027475; AAH27475.1; -.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00041; fn3_1.
DR SMART; SM00060; FN3_1.
DR PROSITE; PS50853; FN3_1.

```



DR PROSITE; PS50835; IG\_LIKE; 1.  
SQ SEQUENCE 324 AA; 33164 MW; 6DC6E0871E227E2A CRC64;

Query Match 51.0%; Score 1700; DB 2; Length 324;  
Best Local Similarity 99.7%; Pred. No. 4.3e-98;

Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 312 MHVGPDDRLVGNSSRAAFPNGLTEIGVTGADAGGYTCIATNPAGEATARVELRVLA 371
    |||||
Db 1 MHVGPDDRLVGNSSRAAFPNGLTEIGVTGADAGGYTCIATNPAGEATARVELRVLA 60
    |||||

QY 372 PHGNSSAEGGRPGPSDIAASARTAAEGSGTLESEPAVOYTEVTATSGLVSWGGRPADP 431
    |||||
Db 61 PHGNSSAEGGRPGPSDIAASARTAAEGSGTLESEPAVOYTEVTATSGLVSWGGRPADP 120
    |||||

QY 432 VMWFQIQNSSSEDETLIYRIVPASSHHFLKHLVPGADYDLCILSPAAGPSDLTATRL 491
    |||||
Db 121 VMWFQIQNSSSEDETLIYRIVPASSHHFLKHLVPGADYDLCILSPAAGPSDLTATRL 180
    |||||

QY 492 LGCARFSTLPASPLCHALQAHVLGTLTVAVGSLVAALLVFTVALLVRGAGNGRLPL 551
    |||||
Db 181 LGCARFSTLPASPLCHALQAHVLGTLTVAVGSLVAALLVFTVALLVRGAGNGRLPL 240
    |||||

QY 552 KLSHVQSQTNGGSPPTPKAHPPRSPPQRSCSLDLDGAGCYGYARRLGAWARRSHSV 611
    |||||
Db 241 KLSHVQSQTNGGSPPTPKAHPPRSPPQRSCSLDLDGAGCYGYARRLGAWARRSHSV 300
    |||||

QY 612 HGGLLGAGCGVGVGSAERLEESYV 635
    |||||
Db 301 HGGLLGAGCGVGVGSAERLEESYV 324
    |||||
```

Search completed: November 17, 2004, 22:40:08  
Job time : 201 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 17, 2004, 22:30:58 / Search time 73 Seconds  
(without alignments)  
3120.453 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336  
Sequence: 1 MAPPLLLILASGAACPLP.....LGAGCRGVGSAERLEESVY 635

Scoring table: BLOSUM62  
Gapop 10.0 , Gapevt 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3331	99.9	635	6	AAO26256 MDT rela
2	3330	99.8	635	5	AAE23980 Human LP2
3	3219	96.5	778	5	ABP70144 Human NOV
4	3218.5	96.5	618	7	AD121104
5	3026.5	90.7	647	5	ABP70142 Human NOV
6	2970	89.0	565	7	ADP08361 Novel pro
7	2773.5	83.1	566	5	ABP70143 Human NOV
8	2773	83.1	551	5	AAE17484 Human leu
9	2568.5	77.0	526	4	ABG04827 Novel hum
10	2475	74.2	468	4	AA870072 Human sec
11	2475	74.2	468	5	ABG65511 Human alb
12	2475	74.2	468	8	ADL78778 Human hno
13	1662.5	49.8	785	3	AA812448 Human bira
14	1662.5	49.8	789	3	AA809968 Human MP5
15	1662.5	49.8	789	7	ADP69106 Human MP5
16	1655.5	49.6	789	4	AA839059 Human pol
17	1655.5	49.6	789	4	AA839059 Human pol
18	1564.5	46.9	636	4	AA832870 Novel hum
19	1557	46.7	628	4	AA832870 Novel hum
20	1556	46.6	628	4	AA832870 Novel hum
21	1556	46.6	628	4	AA832870 Novel hum
22	1556	46.6	628	4	AA832870 Novel hum
23	1556	46.6	628	4	AA832870 Novel hum
24	1556	46.6	628	4	AA832870 Novel hum
25	1556	46.6	628	4	AA832870 Novel hum

26	1546.5	46.4	627	5	ABG34079 Human PRO
27	1546.5	46.4	627	6	ADA01368 Human PRO
28	1546.5	46.4	627	6	ADA43797 Human sec
29	1546.5	46.4	627	6	ADA43565 Human sec
30	1546.5	46.4	627	6	ADA01240 Human PRO
31	1546.5	46.4	627	7	ADA01124 Human sec
32	1546.5	46.4	627	7	ADA43681 Human PRO
33	1546.5	46.4	627	7	ADA06943 Human PRO
34	1546.5	46.4	627	7	ADA08431 Human PRO
35	1546.5	46.4	627	7	ADA09724 Human PRO
36	1546.5	46.4	627	7	ADA09724 Human PRO
37	1546.5	46.4	627	7	ADA09724 Human PRO
38	1546.5	46.4	627	7	ADA09724 Human PRO
39	1546.5	46.4	627	7	ADA09724 Human PRO
40	1546.5	46.4	627	7	ADA09724 Human PRO
41	1546.5	46.4	627	7	ADA09724 Human PRO
42	1546.5	46.4	627	7	ADA09724 Human PRO
43	1546.5	46.4	627	7	ADA09724 Human PRO
44	1546.5	46.4	627	7	ADA09724 Human PRO
45	1546.5	46.4	627	7	ADA09724 Human PRO

## ALIGNMENTS

RESULT 1	AAO26256	standard, protein, 635 AA.
ID	AAO26256	
XX	AAO26256	
AC	AAO26256	
XX	AAO26256	
DT	10-APR-2003	(first entry)
XX	10-APR-2003	
DE	MDDT related human protein SEQ ID No 34.	
XX	MDDT related human protein SEQ ID No 34.	
XX	Cystostatic; antiatherosclerotic; osteopathic; antiarteriosclerotic;	
KW	hepatotropic; antiparasitic; antiallergic; antianaemic; antiaesthetic;	
KW	antithyroid; antinflammatory; antihelminthic; antidiabetic; nephrotropic;	
KW	ophthalmologic; immunosuppressive; dermatological; antifungal;	
KW	antihemematic; antirachitic; antibacterial; vitruic; fungicide;	
KW	antiparasitic; protozoacide; tranquiliser; vulnerary; anti-HIV;	
KW	neurotropic; neuroprotective; anticonvulsant; cerebroprotective;	
KW	neuroleptic; molecules for disease detection and treatment; MDDT;	
KW	immunogen; cancer; actinic keratosis; arteriosclerosis; atherosclerosis;	
KW	burstis; cirrhosis; hepatitis; psoriasis; AIDS; rheumatoid arthritis;	
KW	adult respiratory distress syndrome; Addison's disease; allergy; anaemia;	
KW	asthma; osteoporosis; autoimmune; haemolytic anaemia; scleroderma;	
KW	autoimmune thyroiditis; Crohn's disease; atopic dermatitis;	
KW	diabetes mellitus; Graves' disease; glomerulonephritis;	
KW	systemic lupus erythematosus; systemic sclerosis; ulcerative colitis;	
KW	haemodialysis; uveitis; trauma; Alzheimer's; Pick disease;	
KW	Parkinson disease; amyotrophic lateral sclerosis; epilepsy; stroke;	
KW	Huntington's disease; multiple sclerosis; dementia;	
KW	extrapyramidal disorder; motor neuron disorder; central nervous system;	
KW	neuromuscular disorder; metabolic; endocrine; toxic myopathy;	
XX	periodic paralysis; mental disorder; human.	
OS	Homo sapiens.	
XX	WO200296951-A1.	
PN	05-DEC-2002.	
XX	24-MAY-2002; 2002WO-US016676.	
PF	25-MAY-2001; 2001US-0293723P.	
XX	01-JUN-2001; 2001US-0295257P.	
PR	08-JUN-2001; 2001US-0297220P.	
PR	21-JUN-2001; 2001US-0300526P.	
PR	29-JUN-2001; 2001US-0301874P.	
PR	22-FEB-2002; 2002US-0359413P.	
XX	(INCY-) INCYTE GENOMICS INC.	

XX Tang TY, Yue H, Baughn MR, Dugan BM, Warren BA, Bandman O;  
 PI Richardson TW, Burford N, Sanjamaala B, Becha SP, Yao MG, Yang J;  
 PI Titat UK, Hafalia AUA, Griffin JA, Swatnakar A, Elliott VS;  
 PI Recton SA, Khan FA, Lee EA, Yue H, Lu DM, Maia NK, Thangavelu K;  
 PI Aryazu CS, Xu Y, Ison CH, Huang J, Ding L, Homchell CD;  
 PI Borowsky ML, Emerling BM, Peterson DP, Lu Y, Ramkumar J, Mason PM;  
 PI Zeharjadian Y, Azimzai Y, Stuve LL, Kamigaki LL, Barroso I, Lee S;  
 PI Kable AE;  
 XX  
 DR WPI: 2003-140448/13.  
 DR N-PSDB: AAK93609.  
 XX  
 XX Novel molecules for disease detection and treatment and polynucleotide  
 PT encoding them useful for diagnosing, preventing or treating cell  
 PT proliferative, autoimmune/inflammatory, neurological and developmental  
 PT disorders.  
 XX  
 PS Claim 89, Page 218-220, 260pp, English.

The invention relates to an isolated polypeptide chosen from molecules for disease detection and treatment (MDMT), comprising a one of 39 114-1230 residue amino acid sequences, given in the specification, or a biologically active or immunogenic fragment of the isolated polypeptide. The isolated polypeptide is useful for screening a compound for effectiveness as an agonist or antagonist of the isolated polypeptide. The isolated polypeptide is also useful as an immunogen for preparing polyclonal or monoclonal antibodies by hybridoma technology. The isolated polypeptide and its encoding polynucleotide are useful for diagnosis, treatment and prevention of cancer, actinic keratosis, arteriosclerosis, atherosclerosis, burns, cirrhosis, hepatitis, psoriasis, AIDS, adult respiratory distress syndrome, Addison's disease, allergies, anaemia, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic anaemia, autoimmune thyroiditis, Crohn's disease, atopic dermatitis, diabetes mellitus, Graves' disease, glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections, trauma, Alzheimer's and Pick disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy, stroke, Huntington's disease, multiple sclerosis, dementia, and other extrapyramidal disorder, motor neuron disorder, and other developmental disorders of the central nervous system, neuromuscular disorders, metabolic, endocrine and toxic myopathies, periodic paralyses, mental disorders including mood, anxiety and schizophrenic disorders, anaemia, renal tubular acidosis, epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and cataract. This sequence represents a human MDMT protein relating to the invention

SQ Sequence 635 AA;

Query Match	99.9%	Score 3331	DB 6	Length 635
Best Local Similarity	99.8%	Pred. No. 2e-250		
Matches 634; Conservative	1	Mismatches	0	Gaps 0

QY	1	MAPPELLILLIAGAAACCPICVCONNSEISTSTICARGLLFPDPVNDRTVEIRLADNFI	60
Db	1	MAPPELLILLIAGAAACCPICVCONNSEISTSTICARGLLFPDPVNDRTVEIRLADNFI	60
QY	61	QALGPDPFRNMTGLVDLTLSRNAITRIGARAFGLDESLRSLHDGRLVELGTSLRGV	120
Db	61	QALGPDPDFRSMGTGLVDLTLSRNAITRIGARAFGLDESLRSLHDGRLVELGTSLRGV	120
QY	121	NIQGHILISNQIGRIAPGAFDDPFLSELDLDSYNNLRQVPMAIGAMPALHTLNIDHML	180
Db	121	NIQGHILISNQIGRIAPGAFDDPFLSELDLDSYNNLRQVPMAIGAMPALHTLNIDHML	180
QY	181	IDALPPGAFAPQLGOLSRDLTLSNRNLATLAPDPLFSRGRDAEASPAVLVLSFGSNPLHCNC	240
Db	181	IDALPPGAFAPQLGOLSRDLTLSNRNLATLAPDPLFSRGRDAEASPAVLVLSFGSNPLHCNC	240
QY	241	ELIMLRRLARPPDLDTCASPPGIAGRYFMAVAVEGESCEPPLIARTTORLMTIEGGRATL	300
Db	241	ELIMLRRLARPPDLDTCASPPGIAGRYFMAVAVEGESCEPPLIARTTORLMTIEGGRATL	300

QY	301	RCRALGDPAPFMHWGPPDDRLVGNSSRRAPAFNGTLEICVTASGAGGYTCIATNPAGA	360
Db	301	RCRALGDPAPFMHWGPPDDRLVGNSSRRAPAFNGTLEICVTASGAGGYTCIATNPAGA	360
QY	361	TARVELRYLALPHGNSSAEGGRRPSPDIASAARTAAESEGTLSEFPANQVTEVATTSGL	420
Db	361	TARVELRYLALPHGNSSAEGGRRPSPDIASAARTAAESEGTLSEFPANQVTEVATTSGL	420
QY	421	VSWGGRPADPEVMNQIYNSSSEDETLIYRIYPASSHHFLKHLVPGADYDCLIALSPA	480
Db	421	VSWGGRPADPEVMNQIYNSSSEDETLIYRIYPASSHHFLKHLVPGADYDCLIALSPA	480
QY	481	AGPSLTLTRLGCMHSTLTPASPLCHALQANVLGTLTVANGVYVALLVPTVALLYR	540
Db	481	AGPSLTLTRLGCMHSTLTPASPLCHALQANVLGTLTVANGVYVALLVPTVALLYR	540
QY	541	GRGAGNGRLPLKLSHVQQTNGGSPPTKAPHPRRSPRRQSCSLDLDAGCYGARRL	600
Db	541	GRGAGNGRLPLKLSHVQQTNGGSPPTKAPHPRRSPRRQSCSLDLDAGCYGARRL	600
QY	601	GGAMARRSHSVHGCLLGACRGVGSAPLETSYV	635
Db	601	GGAMARRSHSVHGCLLGACRGVGSAPLETSYV	635

ID	AAE23980	standard; protein; 635 AA
XX		
AC	AAE23980;	
XX		
DT	23-SEP-2002	(first entry)

DE Human LP220 secreted protein.

KM	Human, secreted protein; atherosclerosis; Alzheimer's disease; LP220;
KM	diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KM	rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KM	perfusion injury; atherosclerosis; wound healing; transgenic animal;
KM	gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
KM	chromosome 14q13.
OS	Homo sapiens.

OS Homo sapiens.

Accession	Key	Location/Qualifiers
FT	Peptide	1..16
FT		/label=Signal_peptide
FT	Protein	17..635
FT		/note="Mature human Lp220 secreted protein"

PN WO200226801-A2.

PD 04-APR-2002.

PF 14-SEP-2001; 2001WO-US026026.

PR 28-SEP-2000; 2000US-0236088P.

PA (ELLIL ) LILLY &amp; CO ELLI.

PI Su EW, Wang H;

DR WPI; 2002-471259/50.

XX

PT treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis

PS Claim 8; Page 127-129; 145pp; English.

CC The invention relates to human secreted polypeptides designated LP095,



PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338032P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 PR 08-MAR-2002; 2002US-00093463.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ,  
 PI Boldog FI, Li L, Zetseus BD, Tchierne VT, Gangolli EA, Vermet CAM;  
 PI Posa CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK,  
 PI Vena EZ, Malyankar UM, Anderson DW, Paturajan M, Miller CE;  
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gueve VY, Pochart PF;  
 PI Zhong M;  
 XX  
 XX WPI; 2002-732824/79.  
 DR N-PSDB; ABV99422.  
 XX  
 PT New NOVX polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemia, obesity, immune or hematopoietic  
 PT disorders, and asthma.  
 XX  
 PS Claim 1; Page 264; 619pp; English.  
 XX  
 CC The present invention relates to new isolated proteins (NOVX) and their  
 CC coding sequences (ABV99327-ABV99595 and ABB70049-ABP70149), where X is  
 CC any number from 1 to 48. The NOVX proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The NOVX coding sequences and proteins are useful for treating:  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorders, hematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOVX  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods  
 CC  
 XX  
 SQ Sequence 778 AA;  
 Query Match 96.5%; Score 3219; DB 5; Length 778;  
 Best Local Similarity 96.1%; Pred. No. 1.4e-241;  
 Matches 617; Conservative 2; Mismatches 5; Indels 18; Gaps 1;  
 QY 1 MAPRLILLASGAACPVCVCONLSBSLTLCAHNGLLFVPPVNDRIYELRLADNFI 60  
 DB 1 MAPRLILLASGAACPVCVCONLSBSLTLCAHNGLLFVPPVNDRIYELRLADNFI 60  
 QY 61 QALGPPDFRNMGTGVDTLTSSNATRTIGARAFGLSRSRLHNGNLVEIGTSLGPPV 120  
 DB 61 QALGPPDFRNMGTGVDTLTSSNATRTIGARAFGLSRSRLHNGNLVEIGTSLGPPV 120  
 QY 121 NLGHLISGNQLGRIAGAPDFLESIEDLDSYNNLRQYFWAGIGAMPALHTLNDNL 180  
 DB 121 NLGHLISGNQLGRIAGAPDFLESIEDLDSYNNLRQYFWAGIGAMPALHTLNDNL 180  
 QY 181 IDALPGAPAFQGLSLDLTNSRLATLAPDFLSRGRDASAPAPLVLSGNGPLHCNC 240  
 DB 181 IDALPGAPAFQGLSLDLTNSRLATLAPDFLSRGRDASAPAPLVLSGNGPLHCNC 240

QY 241 ELLMLRLARPDDLETCASPPGCIAGRYFWAVPBGERSCEPRLIARHTQRLWLEGGQATL 300  
 DB 241 ELLMLRLARPDDLETCASPPGCIAGRYFWAVPBGERSCEPRLIARHTQRLWLEGGQATL 300  
 QY 301 RCRLGDPAPPTMHWPPDBRLVGNSSRARAFPGTLEIGVTGAGDAGGYCIAITNPAGEA 360  
 DB 301 RCRLGDPAPPTMHWPPDBRLVGNSSRARAFPGTLEIGVTGAGDAGGYCIAITNPAGEA 360  
 QY 361 TARVELRVIALPHGNSSAEGGR-----PPSDTASARTPAEGGT 402  
 DB 361 TARVELRVIALPHGNSSAEGGRPCPRTSAPMWEQCCGCGPPSDTAAASARTPAEGGT 420  
 QY 403 LEESEPAVQTEVMTATSGLVNMGWRPADPTMWFQIYNNSEDETLIYRIIPASSHFLTK 462  
 DB 421 LEESEPAVQTEVMTATSGLVNMGWRPADPTMWFQIYNNSEDETLIYRIIPASSHFLTK 480  
 QY 463 HLYPGADVDYDCLLALSPAAGPSDLTATRLIGCAHFTLPAAPLCHALQAHVLGTLTAV 522  
 DB 481 HLYPGADVDYDCLLALSPAAGPSDLTATRLIGCAHFTLPAAPLCHALQAHVLGTLTAV 540  
 QY 523 GGVLVAALVFTYALLVYRGRGAGNGRLPLKLSHYQQTNGGSEPTPKAHPRPSPRPOR 582  
 DB 541 GGVLVAALVFTYALLVYRGRGAGNGRLPLKLSHYQQTNGGSEPTPKAHPRPSPRPOR 600  
 QY 583 SCSLDGDAGCYGARPLGGAMARRSHSVHGILGACRGVG 624  
 DB 601 SCSLDGDAGCYGARPLGGAMARRSHSVHGILGACRGVG 642  
 RESULT 4  
 ADI21104  
 ID ADI21104 standard; protein; 618 AA.  
 XX  
 AC ADI21104;  
 XX  
 DT 15-APR-2004 (first entry)  
 XX  
 DE Novel human protein #79.  
 XX  
 XX forensic; nutritional source; damaged tissue; diseased tissue;  
 KM myeloid cell disorder; lymphoid cell disorder;  
 KM bone cartilage tissue growth; tendon tissue growth;  
 KM ligament tissue growth; nerve tissue growth; regeneration; wound healing;  
 KM tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.  
 XX  
 OS Homo sapiens.  
 OS  
 XX WO2003025148-A2.  
 XX  
 PD 27-MAR-2003.  
 XX  
 PF 19-SEP-2002; 2002MO-US029964.  
 XX  
 PR 19-SEP-2001; 2001US-0323739P.  
 PR 13-SEP-2002; 2002US-00323739.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QH, Wang J;  
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou F, Drmanac RT, Wang D;  
 PI Haley-Vicente D;  
 XX  
 DR WPI; 2003-354603/33.  
 DR N-PSDB; ADI21820.  
 XX  
 PT New polynucleotides and secreted proteins, useful for treating myeloid or  
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve  
 PT tissue growth or regeneration, in wound healing, and in tissue repair and  
 PT replacement.  
 XX  
 PS Claim 20; SEQ ID NO 355; 156pp; English.  
 XX



CC The invention relates to an isolated polynucleotide encoding a  
CC polypeptide with biological activity. The polynucleotides and  
CC polypeptides are useful in diagnostics, forensics, gene mapping,  
CC identification of mutations responsible for genetic disorders and other  
CC traits, to assess biodiversity, as nutritional sources or supplements.  
CC The polynucleotides may also be used as molecular weight markers,  
CC chromosome markers or map related gene positions, or as an antigen to  
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are  
CC useful for raising antibodies, as markers for tissues in which the  
CC corresponding polypeptide is expressed, for re-engineering damaged or  
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in  
CC bone cartilage, tendon, ligament and/or nerve tissue growth or  
CC regeneration, in wound healing, in tissue repair and replacement, in  
CC healing of burns, incisions and ulcers, and in treating cancer. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein.  
CC  
XX

SQ Sequence 618 AA;

Query Match 96.5%; Score 3218.5; DB 7; Length 618;  
Best Local Similarity 97.0%; Pred. No. 1.1e-241;  
Matches 616; Conservative 0; Mismatches 2; Indels 17; Gaps 1;

QY 1 MAPPILLILLIAGAAACPLPCVCONTSESLSTCAHRLIFVPPVDRTEIRLADNFI 60  
DB 1 MAPPILLILLIAGAAACPLPCVCONTSESLSTCAHRLIFVPPVDRTEIRLADNFI 60  
QY 61 QALGPPDFRMTGLVLTLSRNAITRIGARFGDLSELSIHMDGNRIYELGTSGIRGPV 120  
DB 61 QALGPPDFRMTGLVLTLSRNAITRIGARFGDLSELSIHMDGNRIYELGTSGIRGPV 120  
QY 121 NIQHLLISGNQGRIPGAFDDFLESLEDLDSYNNLRQVPMAGAMPALHTLNDHL 180  
DB 121 NIQHLLISGNQGRIPGAFDDFLESLEDLDSYNNLRQVPMAGAMPALHTLNDHL 180  
QY 181 IDALPPGAPAOQLGSLRLDITSNRALTAPDPLFSGRDAEASPAVLVSFGNPLHCNC 240  
DB 181 IDALPPGAPAOQLGSLRLDITSNRALTAPDPLFSGRDAEASPAVLVSFGNPLHCNC 240  
QY 241 ELLMLRLARPDDLETCAASPGLAGRYFMAVPEGSECEPPLIARHTORLMTVEGORALT 300  
DB 241 ELLMLRLARPDDLETCAASPGLAGRYFMAVPEGSECEPPLIARHTORLMTVEGORALT 300  
QY 301 RCRAIDDPATMTMVGPDRLVGNSSRABAFPMGTLEIGVTGADAGYTCIATNPAGBA 360  
DB 301 RCRAIDDPATMTMVGPDRLVGNSSRABAFPMGTLEIGVTGADAGYTCIATNPAGBA 360  
QY 361 TARVELRVALLPHGNSABGGRPPSDIAASAKTAAEGGTLESEPAVQVTEVTATSGI 420  
DB 361 TARVELRVALLPHGNSABGGRPPSDIAASAKTAAEGGTLESEPAVQVTEVTATSGI 420  
QY 421 VSWGPRPADPYVMFOIYNSSEDETLIRIVPASHHPLIKHLVPGADYDCLIALSPA 480  
DB 421 VSWGPRPADPYVMFOIYNSSEDETLIRIVPASHHPLIKHLVPGADYDCLIALSPA 480  
QY 481 AGPSDULATRLTGCAHFSTLPASPLCHALQAHVIGSTLVAVGVVLAALLVFTVALVR 540  
DB 481 AGPSDULATRLTGCAHFSTLPASPLCHALQAHVIGSTLVAVGVVLAALLVFTVALVR 540  
QY 541 GRGAGGRPLPLKISHVQSQTNNGSPPTPKAHPRRSPRRPORSCLIDLGAGCYGARRI 600  
DB 541 GRGAGGRPLPLKISHVQSQTNNGSPPTPKAHPRRSPRRPORSCLIDLGAGCYGARRI 600  
QY 601 GGAMARRSHSVHGGLIGAGRGVGSAAERLEESVV 635  
DB 601 GGAMARRSHSVHGGLIGAGRGVGSAAERLEESVV 635  
QY 635 GGAMARRSHSVHGGLIGAGRGVGSAAERLEESVV 618  
DB 635 GGAMARRSHSVHGGLIGAGRGVGSAAERLEESVV 618

RESULT 5  
ID ABP70142  
XX ABP70142 standard; protein; 647 AA.  
AC ABP70142;

XX 27-JAN-2003 (first entry)  
XX  
DE Human NOV44a.  
XX  
XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;  
XX antiinflammatory; cardiatic; haemostatic; neuroprotective; anorectic;  
XX neurotopic; immunosuppressive; osteopathic; antiparkinsonian; cancer;  
XX antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
XX metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
XX immune disorder; haematopoietic disorder; cardiovascular disorder;  
XX bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
XX metabolic syndrome X; wasting disorder; cell differentiation;  
XX cell proliferation; haematopoiesis; wound healing; angiogenesis.  
XX  
OS Homo sapiens.  
XX  
PN WO200272771-A2.  
XX  
PD 19-SEP-2002.  
XX  
PE 08-MAR-2002; 2002MO-US007288.  
XX  
XX 08-MAR-2001; 2001US-0274101P.  
XX 08-MAR-2001; 2001US-0274194P.  
XX 08-MAR-2001; 2001US-0274281P.  
XX 08-MAR-2001; 2001US-0274322P.  
XX 09-MAR-2001; 2001US-0274849P.  
XX 12-MAR-2001; 2001US-0275355P.  
XX 13-MAR-2001; 2001US-0275578P.  
XX 13-MAR-2001; 2001US-0275579P.  
XX 13-MAR-2001; 2001US-0275601P.  
XX 14-MAR-2001; 2001US-0276000P.  
XX 16-MAR-2001; 2001US-0276767P.  
XX 19-MAR-2001; 2001US-0276949P.  
XX 20-MAR-2001; 2001US-0277239P.  
XX 20-MAR-2001; 2001US-0277321P.  
XX 20-MAR-2001; 2001US-0277327P.  
XX 20-MAR-2001; 2001US-0277328P.  
XX 21-MAR-2001; 2001US-0277791P.  
XX 22-MAR-2001; 2001US-0277833P.  
XX 23-MAR-2001; 2001US-0278152P.  
XX 26-MAR-2001; 2001US-0278894P.  
XX 27-MAR-2001; 2001US-0278999P.  
XX 27-MAR-2001; 2001US-0279036P.  
XX 28-MAR-2001; 2001US-0279344P.  
XX 30-MAR-2001; 2001US-0279959P.  
XX 30-MAR-2001; 2001US-0280233P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280822P.  
XX 02-APR-2001; 2001US-0280822P.  
XX 04-APR-2001; 2001US-0281194P.  
XX 13-APR-2001; 2001US-0283675P.  
XX 30-APR-2001; 2001US-0287424P.  
XX 02-MAY-2001; 2001US-0288066P.  
XX 03-MAY-2001; 2001US-0288342P.  
XX 03-MAY-2001; 2001US-0288528P.  
XX 15-MAY-2001; 2001US-0291190P.  
XX 16-MAY-2001; 2001US-0291099P.  
XX 16-MAY-2001; 2001US-0291240P.  
XX 30-MAY-2001; 2001US-0294485P.  
XX 31-MAY-2001; 2001US-0294489P.  
XX 31-MAY-2001; 2001US-0294899P.  
XX 18-JUN-2001; 2001US-0299027P.  
XX 19-JUN-2001; 2001US-0299033P.  
XX 19-JUN-2001; 2001US-0299310P.  
XX 10-JUL-2001; 2001US-0304354P.  
XX 31-JUL-2001; 2001US-0309198P.  
XX 16-AUG-2001; 2001US-0312903P.  
XX 10-SEP-2001; 2001US-0318462P.  
XX 12-SEP-2001; 2001US-0318770P.  
XX 27-SEP-2001; 2001US-0325430P.

PR 27-SEP-2001; 2001US-0325681P.  
 PR 18-OCT-2001; 2001US-0330380P.  
 PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338052P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 PR 08-MAR-2002; 2002US-00093463.  
 PA (CURA-) CURAGEN CORP.  
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ,  
 PI Baidog FL, Li L, Zethusen BD, Tchenev VT, Gangolli EA, Vernet CAM,  
 PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK,  
 PI Voss EZ, Malyankar UM, Anderson DW, Paturajan M, Miller CE,  
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gueev VY, Pochart PF,  
 PI Zhong M;  
 XX WPI; 2002-732824/79.  
 DR N-PSDB; ABV99420.  
 XX  
 PT New NOVX polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
 PT disorders, and asthma.  
 XX  
 PS Claim 1; Page 262; 619pp; English.  
 XX  
 CC The present invention relates to new isolated proteins (NOVX) and their  
 CC coding sequences (ABV99327-ABV99595 and ABV70049-ABV70149), where X is  
 CC any number from 1 to 48. The NOVX proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The NOVX coding sequences and proteins are useful for treating, diabetes,  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorders, hematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOVX  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods  
 XX  
 SQ Sequence 647 AA;  
 XX  
 Query Match 90.7%; Score 3026.5; DB 5; Length 647;  
 Best Local Similarity 93.9%; Pred. No. 1.1e-226;  
 Matches 589; Conservative 0; Mismatches 21; Indels 17; Gaps 3;  
 XX  
 QY 1 MAPPLILLILASGAACPPLPCVCONLSLSITLCARHGLLVFPNVDRRTVELRLADNFI 60  
 DB 1 MAPPLILLILASGAACPPLPCVCONLSLSITLCARHGLLVFPNVDRRTVELRLADNFI 60  
 QY 61 QALGPPFRFMTGIVDTITSRNATTRTGARAFGLSELSRSLHDGNLVELGTSLGSPV 120  
 DB 61 QALGPPFRFMTGIVDTITSRNATTRTGARAFGLSELSRSLHDGNLVELGTSLGSPV 120  
 QY 121 NLQHLISGQOLGRIAGAFDDPLESLEDLSTYNNROYVWAGIGAMPALHTINTLHNH 180  
 DB 121 NLQHLISGQOLGRIAGAFDDPLESLEDLSTYNNROYVWAGIGAMPALHTINTLHNH 180  
 QY 181 IDALPPGAFAQLGQLSLDLTSLNRLATLADPLFRSGRDAEAPPAVLVLSFGNPLHCNC 240

DB 181 IDALPPGAFAQLGQLSLDLTSLNRLATLADPLFRSGRDAEAPPAVLVLSFGNPLHCNC 240  
 QY 241 ELLMLRLARPDDLETCASPGLAGRYFWAVPGEFSCEPPLTARHTORLVLBEGORATL 300  
 DB 241 ELLMLRLARPDDLETCASPGLAGRYFWAVPGEFSCEPPLTARHTORLVLBEGORATL 300  
 QY 301 RCALGDPAPPTMMWVGDDDLVGNSSRARAFPMGTLEIGVTGADGAGYTCIATNPAGEA 360  
 DB 301 RCALGDPAPPTMMWVGDDDLVGNSSRARAFPMGTLEIGVTGADGAGYTCIATNPAGEA 360  
 QY 361 TARVELRVLALPHGNGSSAAGRPGPSDIAASARTAAEGGTLESEBAVQTEVTATSGL 420  
 DB 361 TARVELRVLALPHGNGSSAAGRPGPSDIAASARTAAEGGTLESEBAVQTEVTATSGL 420  
 QY 421 VSWGPRPADPVMFQIQVNSSEDETLIRIVASSHHFLKXVLPGADVDLCIALISPA 480  
 DB 421 VSWGPRPADPVMFQIQVNSSEDETLIRIVASSHHFLKXVLPGADVDLCIALISPA 480  
 QY 481 AGPSDLTATRLGCAHPSSTLPASPLCHALQAHVAGTLTVAVGVVAALLVFTVALLVR 540  
 DB 481 AGPSDLTATRLGCAHPSSTLPASPLCHALQAHVAGTLTVAVGVVAALLVFTVALLVR 540  
 QY 541 GRGAGNGRLPLKLSHVOSQTNQSPSPTPKAHPPRSP-----PRPORSCLDLDAG 592  
 DB 541 GRGAGNGRLPLKLSHVOSQTNQSPSPTPKAHPPRSP-----PRPORSCLDLDAG 592  
 QY 593 CYGYARRLGG----AMARRSHVHG 614  
 DB 593 CQAPGRSLGTEPLCAMGAARGVPG 623  
 XX  
 RESULT 6  
 ADE08361  
 ID ADE08361 standard; protein; 565 AA.  
 XX  
 AC ADE08361;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 XX Novel protein (useful for identifying genetic disorders) #516.  
 DE  
 XX Novel gene; novel protein; tissue marker; molecular weight marker;  
 KW chromosome marker; genetic disorder.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003054152-A2.  
 XX  
 PD 03-JUL-2003.  
 XX  
 PF 10-DEC-2002; 2002WO-US039555.  
 XX  
 PR 10-DEC-2001; 2001US-0339739P.  
 PR 11-DEC-2001; 2001US-0339453P.  
 PR 14-MAR-2002; 2002US-0365091P.  
 PR 14-MAR-2002; 2002US-0365384P.  
 PR 12-APR-2002; 2002US-0372381P.  
 PR 12-APR-2002; 2002US-0372615P.  
 PR 22-APR-2002; 2002US-0012855P.  
 PR 24-APR-2002; 2002US-0376045P.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Auendi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J,  
 PI Ghosh M, Xue AJ, Wehman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
 XX WPI; 2003-569235/53.  
 DR N-PSDB; ADE07450.  
 XX  
 PT New polynucleotides, useful for expressing recombinant proteins for  
 PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

XX Claim 20; SEQ ID NO 1427; 1177pp; English.

XX The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present amino acid sequence represents a protein  
CC of the invention.

XX Sequence 565 AA;

Query Match 89.0%; Score 2970; DB 7; Length 565;  
Best Local Similarity 100.0%; Pred. No. 2.2e-222;  
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 71 MTCGLVLTLSRNAITMIGARAFGLDLSRLSLHDGRLVLELGTSLRGPNLQHLILSGN 130
Db 1 MTCGLVLTLSRNAITMIGARAFGLDLSRLSLHDGRLVLELGTSLRGPNLQHLILSGN 60
QY 131 QUGRIIPGAFDDLELDELDSYNNLRQVPMAGIGAMPALMTLNDHNLIDALPGARA 190
Db 61 QUGRIIPGAFDDLELDELDSYNNLRQVPMAGIGAMPALMTLNDHNLIDALPGARA 120
QY 191 QUGLSRLDLTSNRILTALADPLFSRGRDAEAPLVSFSGNPILHNCCELLMTLRAR 250
Db 121 QUGLSRLDLTSNRILTALADPLFSRGRDAEAPLVSFSGNPILHNCCELLMTLRAR 180
QY 251 PDDLETCASPPGIAGRYPMANVPEGEFSCPEPLIARTTORILWLEGORATLRCALGDPA 310
Db 181 PDDLETCASPPGIAGRYPMANVPEGEFSCPEPLIARTTORILWLEGORATLRCALGDPA 240
QY 311 TMMWGPDDRLVNSSRRARAFPMGTLEIGTGAGDAGCTTCTATNPAGATARVELRVTA 370
Db 241 TMMWGPDDRLVNSSRRARAFPMGTLEIGTGAGDAGCTTCTATNPAGATARVELRVTA 300
QY 371 LPHGNSNAGGRPGPSDIAASARTAEGETLESEPAVOVEVTATSGLVSGPGRPAD 430
Db 301 LPHGNSNAGGRPGPSDIAASARTAEGETLESEPAVOVEVTATSGLVSGPGRPAD 360
QY 431 PVMWFOIQYNSSEDETLIRIVPASHHFLKHLVPGADYDCLIALSPAGPSDLTATR 490
Db 361 PVMWFOIQYNSSEDETLIRIVPASHHFLKHLVPGADYDCLIALSPAGPSDLTATR 420
QY 491 ILGCAHFTLPPASPLCHALQAHVIGSTLVAVGVTVALLVFTVALIVRGAGNGRLP 550
Db 421 ILGCAHFTLPPASPLCHALQAHVIGSTLVAVGVTVALLVFTVALIVRGAGNGRLP 480
QY 551 IKLSHVQQTNGSPSTPKAHPSPRPQRSCSIDLDGACGYARILGGAMARRSHS 610
Db 481 IKLSHVQQTNGSPSTPKAHPSPRPQRSCSIDLDGACGYARILGGAMARRSHS 540
QY 611 VHGGLIGAGCRGVGSAERLESVV 635
Db 541 VHGGLIGAGCRGVGSAERLESVV 565
```

RESULT 7  
ABP70143  
ID ABP70143 standard; protein; 566 AA.

XX ABP70143;

DT 27-JAN-2003 (first entry)

DE Human NOV44b.

XX Human; anti-HIV; cytostatic; antidiabetic; antiaslomatic; cachexia; AIDS;  
KM antiinflammatory; cardiatic; haemostatic; neuroprotective; anorectic;  
KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;

KW antiinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
KM metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
KM neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
KM immune disorder; hematopoietic disorder; cardiovascular disorder;  
KM bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
KM metabolic syndrome X; wasting disorder; cell differentiation;  
KM cell proliferation; hematopoiesis; wound healing; angiogenesis.  
OS Homo sapiens.  
PN WO200272771-A2.  
XX 19-SEP-2002.  
PD 08-MAR-2002; 2002WC-US007288.  
XX 08-MAR-2001; 2001US-0274101P.  
XX 08-MAR-2001; 2001US-0274194P.  
XX 08-MAR-2001; 2001US-0274281P.  
XX 08-MAR-2001; 2001US-0274322P.  
XX 09-MAR-2001; 2001US-0274849P.  
XX 12-MAR-2001; 2001US-0275235P.  
XX 13-MAR-2001; 2001US-0275578P.  
XX 13-MAR-2001; 2001US-0275579P.  
XX 13-MAR-2001; 2001US-0275601P.  
XX 14-MAR-2001; 2001US-0276000P.  
XX 16-MAR-2001; 2001US-0276776P.  
XX 19-MAR-2001; 2001US-0276994P.  
XX 20-MAR-2001; 2001US-0277239P.  
XX 20-MAR-2001; 2001US-0277321P.  
XX 20-MAR-2001; 2001US-0277327P.  
XX 20-MAR-2001; 2001US-0277381P.  
XX 21-MAR-2001; 2001US-0277791P.  
XX 22-MAR-2001; 2001US-0277833P.  
XX 23-MAR-2001; 2001US-0278152P.  
XX 26-MAR-2001; 2001US-0278894P.  
XX 27-MAR-2001; 2001US-0278999P.  
XX 27-MAR-2001; 2001US-0279036P.  
XX 28-MAR-2001; 2001US-0279344P.  
XX 30-MAR-2001; 2001US-0279995P.  
XX 30-MAR-2001; 2001US-0280233P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280822P.  
XX 02-APR-2001; 2001US-0280900P.  
XX 04-APR-2001; 2001US-0281194P.  
XX 13-APR-2001; 2001US-0283675P.  
XX 30-APR-2001; 2001US-0287424P.  
XX 02-MAY-2001; 2001US-0288066P.  
XX 03-MAY-2001; 2001US-0288342P.  
XX 03-MAY-2001; 2001US-0288528P.  
XX 15-MAY-2001; 2001US-0291190P.  
XX 16-MAY-2001; 2001US-0291099P.  
XX 16-MAY-2001; 2001US-0291240P.  
XX 30-MAY-2001; 2001US-0294485P.  
XX 31-MAY-2001; 2001US-0294889P.  
XX 31-MAY-2001; 2001US-0294899P.  
XX 18-JUN-2001; 2001US-0299027P.  
XX 19-JUN-2001; 2001US-0299303P.  
XX 19-JUN-2001; 2001US-0299310P.  
XX 10-JUL-2001; 2001US-0304354P.  
XX 31-JUL-2001; 2001US-0309198P.  
XX 16-AUG-2001; 2001US-0312903P.  
XX 10-SEP-2001; 2001US-0318462P.  
XX 12-SEP-2001; 2001US-0318770P.  
XX 27-SEP-2001; 2001US-0325430P.  
XX 27-SEP-2001; 2001US-0325681P.  
XX 18-OCT-2001; 2001US-0330380P.  
XX 31-OCT-2001; 2001US-0335301P.  
XX 14-NOV-2001; 2001US-0332172P.  
XX 14-NOV-2001; 2001US-0332271P.  
XX 14-NOV-2001; 2001US-0332272P.  
XX 14-NOV-2001; 2001US-0333184P.  
XX 14-NOV-2001; 2001US-0333272P.



CC adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukaemia,  
 CC kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,  
 CC colon adenocarcinoma, genitourinary tract transitional cell tumours,  
 CC rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas and  
 CC prostate adenocarcinoma. Zlir protein is useful for identifying agonists  
 CC and antagonists of the polypeptide, for drug design, to screen for cell  
 CC metabolism affecting receptors, for analysis of cell phenotype, and as  
 CC animal feed supplement and cell culture components. Zlir DNA is also  
 CC useful in gene therapy. The present sequence is human Zlir8 protein.  
 CC Zlir8 gene is located on chromosome 11q13  
 CC  
 XX

Sequence 551 AA:

Query Match 83.1%; Score 2773; DB 5; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 4,8e-207;  
 Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPPILLILLASGAACPLPCVCONLSLSTLCARHGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCONLSLSTLCARHGLFVPPNVDRTVELRLADNFI 60  
 QY QALGPPDFRMTGAVDLTISRNAITRIGARAFGDLSELRSLHDGNRLVELGSGRGPV 120  
 DB QALGPPDFRMTGAVDLTISRNAITRIGARAFGDLSELRSLHDGNRLVELGSGRGPV 120  
 QY 121 NIOHLILSGNOLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 180  
 DB 121 NIOHLILSGNOLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 180  
 QY 181 IDLPPGAFQOLGOLSRDLTISRNLATLADPLFSRGRDAEAPPLVLSFGNPJHCNC 240  
 DB 181 IDLPPGAFQOLGOLSRDLTISRNLATLADPLFSRGRDAEAPPLVLSFGNPJHCNC 240  
 QY 241 ELIMRLRLARPDDLETCAASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWVLEGORATL 300  
 DB 241 ELIMRLRLARPDDLETCAASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWVLEGORATL 300  
 QY 301 ELMMLRLARPDDLETCAASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWVLEGORATL 300  
 DB 301 ELMMLRLARPDDLETCAASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWVLEGORATL 300  
 QY 301 RCALGDPAFTMHWGPDRLVGNSSRARAFPNGLTLEIGVTGADAGGYTCIATNPAGEA 360  
 DB 301 RCALGDPAFTMHWGPDRLVGNSSRARAFPNGLTLEIGVTGADAGGYTCIATNPAGEA 360  
 QY 361 TARVERLRLALPHGNSAAGGPPSPDIAASARTAAEGGTLESEBPAVQVTEVATSGI 420  
 DB 361 TARVERLRLALPHGNSAAGGPPSPDIAASARTAAEGGTLESEBPAVQVTEVATSGI 420  
 QY 421 VSWGPPRPADPVMMFOIQYNSSEDETLIRIVPASSHHFLKHLVPGADVDLCLALSPA 480  
 DB 421 VSWGPPRPADPVMMFOIQYNSSEDETLIRIVPASSHHFLKHLVPGADVDLCLALSPA 480  
 QY 481 AGPSDLTATRLTGCAHFSTLPASPLCHALQAHVLTGTLTVAVGVVAAL 530  
 DB 481 AGPSDLTATRLTGCAHFSTLPASPLCHALQAHVLTGTLTVAVGVVAAL 530

RESULT 9

ABG04827  
 ID ABG04827 standard; protein; 526 AA.

AC ABG04827;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #4818.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW Food supplement; medical imaging; diagnostic; genetic disorder.  
 XX

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PI Dermanac RT, Liu C, Tang YT;

XX WPI, 2001-639362/73.

DR N-PDB; AAS69014.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 35186; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (I) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

XX useful for generating antibodies against it, detecting or quantitating a

XX polypeptide in tissue, as molecular weight markers and as a food

XX supplement. (II) and its binding partners are useful in medical imaging

XX of sites expressing (II). (I) and (II) are useful for treating disorders

XX involving aberrant protein expression or biological activity. The

XX polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic

XX patent did not appear in the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 526 AA;

Query Match 77.0%; Score 2568.5; DB 4; Length 526;

Best Local Similarity 89.3%; Pred. No. 3.8e-191;

Matches 502; Conservative 4; Mismatches 15; Indels 41; Gaps 2;

QY 1 MAPPILLILLASGAACPLPCVCONLSLSTLCARHGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCONLSLSTLCARHGLFVPPNVDRTVELRLADNFI 60  
 QY QALGPPDFRMTGAVDLTISRNAITRIGARAFGDLSELRSLHDGNRLVELGSGRGPV 119  
 DB QALGPPDFRMTGAVDLTISRNAITRIGARAFGDLSELRSLHDGNRLVELGSGRGPV 119  
 QY 120 NIOHLILSGNOLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 179  
 DB 120 NIOHLILSGNOLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 180  
 QY 181 IDLPPGAFQOLGOLSRDLTISRNLATLADPLFSRGRDAEAPPLVLSFGNPJHCNC 240  
 DB 181 IDLPPGAFQOLGOLSRDLTISRNLATLADPLFSRGRDAEAPPLVLSFGNPJHCNC 240  
 QY 240 CELLMRLRLARPDDLETCAASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWVLEGORAT 299  
 DB 240 CELLMRLRLARPDDLETCAASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWVLEGORAT 300  
 QY 300 IRCRALGDPAFTMHWGPDRLVGNSSRARAFPNGLTLEIGVTGADAGGYTCIATNPAGE 359  
 DB 300 IRCRALGDPAFTMHWGPDRLVGNSSRARAFPNGLTLEIGVTGADAGGYTCIATNPAGE 360  
 QY 360 ATARVERLRLALPHGNSAAGGPPSPDIAASARTAAEGGTLESEBPAVQVTEVATSGI 419  
 DB 360 ATARVERLRLALPHGNSAAGGPPSPDIAASARTAAEGGTLESEBPAVQVTEVATSGI 419





PI Rosen CA, Haseltine WA;  
 XX WPI; 2002-010886/01.  
 DR New fusion protein for treating disease e.g. diabetes comprises an  
 XX albumin fused to a therapeutic protein.  
 PT  
 XX  
 PS Claim 1; Page 2080-2081; 2102pp; English.  
 CC The present invention relates to albumin fusion proteins comprising a  
 CC therapeutic protein X and human albumin (HA, also known as human serum  
 CC albumin, HSA). The proteins are useful for treating a disease or disorder  
 CC that may be modulated by therapeutic protein X. The albumin extends the  
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo  
 CC activity. The protein X is useful for treating and diagnosing disorders  
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's  
 CC disease, ulcerative colitis), immune disorders (e.g. acquired  
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),  
 CC hematopoietic disorders, neural disorders (e.g. Alzheimer's,  
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,  
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).  
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention  
 CC  
 XX  
 SQ Sequence 468 AA;

Query Match 74.2%; Score 2475; DB 5; Length 468;  
 Best Local Similarity 99.8%; Pred. No. 6.3e-184;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPALHTLNDHNIDALPPGAFAQGLGSLRLDTSNRLATLADPILFSGRDAEAPAL 227  
 DB 1 MRALHTLNDHNIDALPPGAFAQGLGSLRLDTSNRLATLADPILFSGRDAEAPAL 60  
 QY 228 VLSFGSNPLHNCCELLMLRLARPDDLTCASPPGAGRYFMAVPEGECSCEPPLARHT 287  
 DB 61 VLSFGSNPLHNCCELLMLRLARPDDLTCASPPGAGRYFMAVPEGECSCEPPLARHT 120  
 QY 288 QULWTEGQRATLRCAAGDPAPYVMWVGPDDRLVGNSSRAAPFPNGTLEIGVAGADAG 347  
 DB 121 QRLWTEGQRATLRCAAGDPAPYVMWVGPDDRLVGNSSRAAPFPNGTLEIGVAGADAG 180  
 QY 348 GTCTCTANPAGEATAVEIRVIALPPGNSAEGGPGSGSDTAAKRTAAEGEGTLESEP 407  
 DB 181 GTCTCTANPAGEATAVEIRVIALPPGNSAEGGPGSGSDTAAKRTAAEGEGTLESEP 240  
 QY 408 AVQVTEVTATSGVSWGPRPADPVWPFQIQYNSSEDETLIRIYVAPASHHFLKHLVPG 467  
 DB 241 AVQVTEVTATSGVSWGPRPADPVWPFQIQYNSSEDETLIRIYVAPASHHFLKHLVPG 300  
 QY 468 AAYDICTLALSPAGPSDLTATRLGCAHFSTLPASPLCHALQAHVLTGTLTVAVGVLY 527  
 DB 301 AAYDICTLALSPAGPSDLTATRLGCAHFSTLPASPLCHALQAHVLTGTLTVAVGVLY 360  
 QY 528 AALLVETVALVRGAGNGRLPKLSHVQSTNGSPSTPRKAPRPPRPQSCSID 587  
 DB 361 AALLVETVALVRGAGNGRLPKLSHVQSTNGSPSTPRKAPRPPRPQSCSID 420  
 QY 588 LGDACCYGYARRLGAMARRSHSVHGLGACRCGVGSAEPLTESV 635  
 DB 421 LGDACCYGYARRLGAMARRSHSVHGLGACRCGVGSAEPLTESV 468

RESULT 12  
 ADL78778  
 ID ADL78778 standard; protein; 468 AA.  
 AC ADL78778;  
 XX  
 XX 20-MAY-2004 (first entry)  
 DE Albumin fusion protein related therapeutic protein X, SEQ ID NO 2260.  
 XX albumin fusion protein; cytoskeletal; antinaemic; antiarthritic;

KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;  
 KW antipsoriatic; antibacterial; osteopathic; dermatological; antigout;  
 KW immunomodulator; antiarrhythmic; cardiac; neurologic; antidiabetic;  
 KW nephrotoxic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;  
 KW antidiabetic; anabolic; hypertensive; vulnery; gene therapy; cancer;  
 KW reproductive system disorder; therapeutic protein.  
 XX  
 OS Unidentified.  
 XX  
 PN US2004010134-A1.  
 PD 15-JAN-2004.  
 XX  
 PF 12-APR-2001; 2001US-00833245.  
 XX  
 PR 12-APR-2000; 2000US-0229358P.  
 PR 25-APR-2000; 2000US-019384P.  
 PR 21-DEC-2000; 2000US-0256931P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (HASE/) HASELTINE W A.  
 XX  
 PI Rosen CA, Haseltine WA;  
 XX  
 DR WPI; 2004-090519/09.  
 XX  
 PT New albumin fusion proteins, useful for diagnosing, treating, preventing  
 PT or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,  
 PT asthma, inflammatory bowel disease or Alzheimer's disease.  
 XX  
 XX Disclosure; SEQ ID NO 2260; 279pp; English.

The invention relates to a novel albumin fusion protein. The invention  
 further relates to: a composition comprising the albumin fusion protein  
 and a pharmaceutical carrier; a kit comprising the albumin fusion protein  
 and albumin fusion protein formula; a method of treating a disease or  
 disorder in a patient comprising the step of administering the albumin  
 fusion protein; a method of treating a patient with a disease or disorder  
 that is modulated by Therapeutic protein X, or its fragment or variant;  
 a method of extending the shelf life of Therapeutic protein X, or its  
 fragment or variant; a nucleic acid molecule comprising a polynucleotide  
 sequence encoding the albumin fusion protein; a vector comprising the  
 nucleic acid molecule of the albumin fusion protein; and a host cell  
 comprising the nucleic acid molecule of the albumin fusion protein. The  
 albumin fusion protein and its compositions have the following  
 activities: cytostatic, antinaemic, antiarthritic, antiasthmatic, anti-  
 HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial,  
 osteopathic, dermatological, antigout, immunomodulator, antiarrhythmic,  
 cardiac, neurotoxic, antiparkinsonian, tranquilizer, antidiabetic, anabolic,  
 neuroprotective, and vulnery. The albumin fusion protein nucleic acid may  
 be used in gene therapy to treat disorders. The albumin fusion protein is  
 useful for diagnosing, treating, preventing or ameliorating diseases or  
 disorders comprising indication: Y. The diseases or disorders include:  
 cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer),  
 immune or haematopoietic diseases (e.g. anemia, Hodgkin's disease, acute  
 lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,  
 autoimmune disease, inflammatory bowel disease, psoriasis or Lyme  
 disease), reproductive system disorders (e.g. prostaticitis, inguinal  
 hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-  
 Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,  
 Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy  
 or cachexia), cardiovascular disease (e.g. thalomyomas, heart disease,  
 arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or  
 hyponatremia), mixed foetal diseases (e.g. foetal alcohol syndrome,  
 Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay-  
 Sachs disease), excretory diseases (e.g. urinary incontinence, urinary  
 tract infections or renal disorders), neural or sensory disease (e.g.  
 Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,  
 cerebellar ataxia, attention deficit disorder, autism or obsessive  
 compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or  
 occupational lung disease), endocrine diseases (e.g. diabetes, Addison's  
 disease or glomerulonephritis), digestive diseases (e.g. portal

CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)  
CC or connective tissue or epithelial diseases (e.g. Crohn's disease,  
CC scleroderma, wound healing or epidermolysis bullosa). This sequence  
CC represents a therapeutic protein X relating to the albumin fusion protein  
CC of the invention. The sequence listing data for this specification was  
CC downloaded from the USPTO website.

XX Sequence 468 AA:

Query Match 74.2%; Score 2475; DB 8; Length 468;

Best Local Similarity 99.8%; Pred. No. 6.3e-184;  
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPRLHTLNDHNLIDLPKPAFQQLGSLDITSLRLATLADPLPSRGDAEASAPL 227  
DB 1 MPRLHTLNDHNLIDLPKPAFQQLGSLDITSLRLATLADPLPSRGDAEASAPL 60  
QY 228 VLSFSGNPLHNCNCELMLRLARPPDLETGASPPGLAGRYFMAVPGEFSCPEPLIARHT 287  
DB 61 VLSFSGNPLHNCNCELMLRLARPPDLETGASPPGLAGRYFMAVPGEFSCPEPLIARHT 120  
QY 288 QRLMWLEGGQATLRCRALGDPATMHWVGPDDRLVGNSSPARAFPNGTLEIGVTGAGDAG 347  
DB 121 QRLMWLEGGQATLRCRALGDPATMHWVGPDDRLVGNSSPARAFPNGTLEIGATGAGDAG 180  
QY 348 GYTCTATNPGEATARVELAVLALPHGNSAAGGPPGSDIAASARTAEGETLESEP 407  
DB 181 GYTCTATNPGEATARVELAVLALPHGNSAAGGPPGSDIAASARTAEGETLESEP 240  
QY 408 AVQVTEVTATSGLVSWGPGRPADPVMMFOIQYNSEDETLIYRIIVPASHHFLKLHPG 467  
DB 241 AVQVTEVTATSGLVSWGPGRPADPVMMFOIQYNSEDETLIYRIIVPASHHFLKLHPG 300  
QY 468 ADVDLCLIALSPAAGPSDLTATRLIGCAHPSLTPASPLCHALQAHVIGTLTVAVGVLV 527  
DB 301 ADVDLCLIALSPAAGPSDLTATRLIGCAHPSLTPASPLCHALQAHVIGTLTVAVGVLV 360  
QY 528 AALLVFTVALIVGRGNGRLPKLSHNSQTMGSPPTPKAHPPSPPPRPORSCSLD 567  
DB 361 AALLVFTVALIVGRGNGRLPKLSHNSQTMGSPPTPKAHPPSPPPRPORSCSLD 420  
QY 588 LGDAGCGYARLRGAMARRSHSVHGGILGACGKGVGSAERLEESYV 635  
DB 421 LGDAGCGYARLRGAMARRSHSVHGGILGACGKGVGSAERLEESYV 468

RESULT 13  
AAB12448  
ID AAB12448 standard; protein; 785 AA.

XX AC AAB12448;

XX 19-OCT-2000 (first entry)

DE Human hh00149 protein SEQ ID NO:4.  
KW Human; ubiquitin-like protein; 149Y2H#151; hh00149; brain;  
KW 2-hybrid screening; neuroprotective; signal transducer;  
KW nervous system disease; diagnosis.

OS Homo sapiens.

XX MO2000031255-A1.

XX 02-JUN-2000.

XX 18-NOV-1999; 99WO-JP006448.

XX 20-NOV-1998; 98JP-0031701.

XX \* (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX Funahashi S, Miyata S;

XX MPI, 2000-400066/34.  
DR N-PSDB; AAB60605.  
XX Gene encoding ubiquitin-like protein which interacts with protein  
PT hh00149, useful in the diagnosis and treatment of diseases associated  
PT with the nervous system.

XX Example 2; Page 74-79; 88pp; Japanese.

PS The present invention describes a ubiquitin-like protein, designated  
CC 149Y2H#151, which interacts with protein hh00149 expressed specifically  
CC in the brain. The 149Y2H#151 protein has neuroprotective activity, and is  
CC a signal transducer. The 149Y2H#151 gene and encoded protein are useful  
CC in the diagnosis and treatment of diseases associated with the nervous  
CC system. The protein can interact with protein hh00149 expressed  
CC specifically in the brain. The present sequence represents the human  
CC hh00149 protein

XX Sequence 785 AA:

Query Match 49.8%; Score 1662.5; DB 3; Length 785;

Best Local Similarity 46.4%; Pred. No. 2.5e-120;  
Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

QY 9 LLASGAA-----ACPLPCVQNLSESLSTLCARGLLPFPVNDRTVELRLADNFIQAL 63  
DB 4 LLAFAFAVADACPKYCCVQNLSESLSTLCARGLLPFPVNDRTVELRLADNFIQAL 63  
QY 64 GPPDFPNTMTGLVDLTLSRNAITRIGARFGLLESLRLHLDGNRLVELGTSLRGVNTQ 123  
DB 64 SRQDFPNTMTGLVDLTLSRNTISHIQFFSLDESLRLSHLDNRLPSLIGDTLRGVNTQ 123  
QY 124 HLLSGNQLGRIAPGAFDFLSELDLDSYNNLRQVPAIGAMALHTLNDHNLIDA 183  
DB 124 HLLVNNNOGLADAEFEFLTLEDLDLSYNNLHGLPMDSVRRVNNLHQLSDHNLIDH 183  
QY 184 LPFGAFQQLGSLDITSLRLATLAPDPLFSRGDA--EASD--APLVLSFGNPLHNCN 239  
DB 184 IAGTADLDLKLRLDITSLRLATLAPDPLFSRGDA--EASD--APLVLSFGNPLHNCN 243  
QY 240 CELLMRLRLARPPDLETGASPPGLAGRYFMAVPGEFSCPEPLIARHTQRLMWLEGGQAT 299  
DB 244 CELLMRLRLARPPDLETGASPPGLAGRYFMAVPGEFSCPEPLIARHTQRLMWLEGGQAT 303  
QY 300 LRCRALGDPATMHWVGPDDRLVGNSSPARAFPNGTLEIGVTGAGDAGTCTATNPAGE 359  
DB 304 LKCKAIGDPSPLTHWAPDDRLVGNSSRTIAYDNGTIDITTSQDSGAFCTIAANAGE 363  
QY 360 ATARVELRLALPHGNSAAGGPPGSDIAASARTAEGETLESEP-----AVQ 410  
DB 364 ATARVELRLALPHGNSAAGGPPGSDIAASARTAEGETLESEP-----AVQ 423  
QY 411 VTEVTATSGLVSWGPGRPADPVMMFOIQYNSEDETLIYRIIVPASHHFLKLHPGADY 470  
DB 424 VSEVTTTSAIVKWSKGAAPRVKMYQLQYNCSDEVLIYRMIIPASNFAVNNNIVSGGY 483  
QY 471 DLCLIALSPAAGPSDLTATRLIGCAHPSLTPASPLCHALQAHVIGTLTVAVGVLVAA 530  
DB 484 DLCLIALSPAAGPSDLTATRLIGCAHPSLTPASPLCHALQAHVIGTLTVAVGVLVAA 542  
QY 531 LVFTVALIVGRGNGRLPKLSHNSQTMGSPPTPKAHPPSPPPRPORSCSLD 577  
DB 543 LVFTVALIVGRGNGRLPKLSHNSQTMGSPPTPKAHPPSPPPRPORSCSLD 601  
QY 578 -----PRP----- 580  
DB 602 NELLDFTASLARSDDSSSSSLGSGAAGLGRAPWRIPSPAPRPKPSLDRLMGAFASIDL 661  
QY 581 -----ORSCSLD 587  
DB 662 KSORKEBLDSRTPAGGACTSARGHHSREPLIGPPAARSLPLPLBCKAKRSHSFD 721

```

QY 588 LGD-----AGCYGARLUGAMARRSHSHVHGILG-----AGCRGVGSABERLE 631
DB 722 MGFFAAAGVAVGGYSPPRKVSINWTKRSLSYNGMLLPFEESDLYGAGGTGSSSEWVM 781
QY 632 ESYV 635
DB 782 ESTV 785

RESULT 14
AAB09968
ID AAB09968 standard; protein; 789 AA.
AC AAB09968;
XX 19-OCT-2000 (first entry)
DE Human brain-specific transmembrane glycoprotein.
XX
XX Transmembrane glycoprotein; human; brain; cellular signal transducer;
XX neuroactive; neuroprotective; cerebroprotective; drug development;
XX treatment; nervous disease; diagnostic.
XX
XX Homo sapiens.
XX WO200031256-A1.
XX 02-JUN-2000.
XX
XX 18-NOV-1999; 99WO-JP006449.
XX
XX 20-NOV-1998; 98BP-00331727.
XX
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX (KAZU-) KAZUSA DNA RES INST.
XX
XX Funahashi S, Miyata S, Nomura N, Nagase T, Ohara O;
XX
XX MPI; 2000-411954/35.
XX N-PSDB; AAA40083.
XX
XX Gene encoding a brain-specific transmembrane glycoprotein that has a
XX typical PDZ protein binding motif and functions as a cellular signal
XX transducer, useful in developing drugs for treating nervous diseases.
XX
XX Claim 1a; Page 52-57; 63pp; Japanese.
XX
XX This invention describes a novel gene encoding a human brain-specific
XX transmembrane glycoprotein that has a typical PDZ protein binding motif
XX and functions as a cellular signal transducer. The product of the
XX invention has neuroactive, neuroprotective and cerebroprotective
XX activity. The gene and encoded protein are useful in developing drugs for
XX treating nervous diseases, and also for studying functions of the nervous
XX system or onset mechanism of nerve-related diseases. They may also be
XX used for the production of diagnostic reagents such as oligonucleotide
XX probes and antibodies for detecting proteins comprising PDZ domains and
XX the nucleic acids that encode them. They may also be used to produce and
XX identify modulators of the proteins expression and activity such as
XX antisense sequences and antibodies. This sequence represents the human
XX brain-specific transmembrane glycoprotein described in the method of the
XX invention
XX
XX Sequence 789 AA;
XX
XX Query Match 49.8%; Score 1662.5; DB 3; Length 789;
XX Best Local Similarity 46.4%; Pred. No.2.5e-120;
XX Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;
QY 9 LIAAGAA-----ACPLPCVCONLSBSISLTLCAHGLLFVPPNVDRTVELRLADNFIAL 63
DB 8 LIAFGMAFAVVDACPFCVCQNLSESLGLTCLPSKGLLFPPDIDRTVELRLGNFIHI 67
QY 64 GPPDFNMNTGLVDLTLSRAIITRIGARAFGDLESLSLHLDGNRLVELGTSLRGPVNIQ 123

```

```

DB 68 SRQDFANMTGLVDLTLSRNTISHIQFSPFLDESLASLHSDNRILPSLSDTLRGLVNIQ 127
QY 124 HLISGNGOIGRIAPGAFDDFLESLEDLDDSYNNLRPOVPAIGAMPALHTLNDHILDA 163
DB 128 HLIVNNNOIGGIADAEFAFEDFLTLLEDLDDSYNNLHGLPMDSVRRMNLHQLSDHNLDDH 187
QY 184 LPPGAFAOGLSRDLDTSNRLATTAPDPLFSGRDA--EASP--APVLVSPGNPLHCN 239
DB 188 IAEGTADLOKLARLDLTNSRLQKPPDPLPASQASALTATFAPPLSFSGGNPLHCN 247
QY 240 CELLMRLRLAPDDULETCASPPGLAGRYFWAVPECFSECEPLIARHTORLWLEGCORAT 299
DB 248 CELLMRLRLERDDULETCSPGLKGRYFWMHREEFVECEPLIHTIKLVLBEQAAAT 307
QY 300 LKCRALGDPAPMTWVGPDDRLVGNSSRRAPAPNGTLEIGVTAGDAGGYTCIATNPAGE 359
DB 308 LKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVVONGTLIDIFITTSQSGAFCTAANAAGB 367
QY 360 ATARVELRLVALPHGNSSAEGGRPGP--SDIAASRTAEGGTLESER-----AYQ 410
DB 368 ATAMVEVSTIVLPHLNSNSTSRTPPKRSLSDLTGSSKTSRGSGSGGSEPPKSPPRAVL 427
QY 411 VTEVTATSGLVSWGPGRPADPYWMFQIQYNSEDEDTLIYRIYVAPSSHHEFLKLHVPADY 470
DB 428 VSEVTTTSLVWVSVKSAAPRYVMYQLOQNCSDDEVLIYRMTPASKAKAFVNNLVSGTGY 487
QY 471 DICTLALSPAAGBSDLTATRLIGCAHFTLPSAPLCAQAHVLTGTTLVAVGVLYAAL 530
DB 488 DICTLALMWDOTA--TTLTATNIVGCAQFFKADYPOQSNHSQLTGTMILVIGIIVALT 546
QY 531 LVFTVALLVGRGAGNGRLPLKL-----SHVOSTNGS-----PSPTKAKHPRSP----- 577
DB 547 LVFTVILMRYK--VCNHEAPSKMAAIVSNVSTQINAQPPPSAPAGAPQGPVKVYR 605
QY 578 -----PRP----- 580
DB 606 NELDFTASLARASDSSSSSSSGCEAGLGRAPWRIPSPAPRPKSLRLIMGAFASLDL 665
QY 581 -----ORCSLSD 587
DB 666 KQCRKEHLDSRTPAAGCAGTSARGHSDREPLGPAPARASLLPLPLEGKAKRSHSPD 725
QY 588 LGD-----AGCYGARLUGAMARRSHSHVHGILG-----AGCRGVGSABERLE 631
DB 726 MGFFAAAGVAVGGYSPPRKVSINWTKRSLSYNGMLLPFEESDLYGAGGTGSSSEWVM 785
QY 632 ESYV 635
DB 786 ESTV 789

RESULT 15
ADf69106
ID ADf69106 standard; protein; 789 AA.
XX
XX ADF69106;
AC
XX
XX 12-FEB-2004 (first entry)
DE Human MP53 protein sequence SEQ ID NO:76.
XX
XX p53 pathway modulating agent; MP53; p53 modulator; cytostatic;
XX gene therapy; cancer; human.
XX
XX Homo sapiens.
XX
XX WO2003083047-A2.
XX
XX 09-OCT-2003.
XX
XX 28-FEB-2003; 2003WO-US006025.
XX

```

PR 01-MAR-2002; 2002US-0361136P.  
 XX (EXEL-) EXELIXIS INC.  
 PA Belvin M, Francis-Lang H, Friedman L, Plozman GD., Heuer TS, Li D;  
 PI Funke RP;  
 XX WPI: 2003-812540/76.  
 DR N-PSDB; ADF69162.  
 XX  
 PT Identifying a candidate p53 pathway modulating agent for treating e.g.,  
 PT cancer by contacting an assay system comprising a MP53 polypeptide or  
 PT nucleic acid with a test agent and detecting a test agent-biased  
 PT activity.  
 PS  
 PS Example; SEQ ID NO 76; 406bp; English.  
 XX  
 CC The present invention describes a method for identifying a candidate p53  
 CC pathway modulating agent, which comprises: (a) providing an assay system  
 CC comprising a MP53 (modulator of p53) polypeptide or nucleic acid or its  
 CC fragment or derivative; (b) contacting the assay system with a test agent  
 CC under conditions where the system provides a reference activity except in  
 CC the presence of the test agent; and (c) detecting a test agent-biased  
 CC activity, where a difference between the test agent-biased activity and  
 CC the reference activity identifies the test agent as a candidate p53  
 CC pathway modulating agent. Also described: (1) modulating the p53 pathway  
 CC of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3)  
 CC diagnosing a disease in a patient. MP53 has cytosolic activity, and can  
 CC be used in gene therapy. The method is useful for identifying a candidate  
 CC p53 pathway modulating agent for preparing a composition for diagnosing  
 CC or treating e.g., cancer. The present sequence represents a human MP53  
 CC protein, which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 789 AA;

Query Match 49.8%; Score 1662.5; DB 7; Length 789;  
 Best Local Similarity 46.4%; Pred. No. 2.5e-120;  
 Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

QY 9 LLASGAA-----ACPLCVCNLSSESLTCAHKGILFVPPNVDRRTVELRLADNFIQAL 63  
 DB 8 LLAFGMAFAVVDACPKYCVCONLSBSLGTLCPSKGLFVPPDIDRTVELRLGNGFIHI 67  
 QY 64 GPPDFRMTGTVDLTLSRNATIRIGARAFGDLRSIHLGDNRLVEIGTSLRGPNVLO 123  
 DB 68 SRQDFAMTGTVDLTLSRNTSHIQPSFLDLSRLSLHDSNRKLPBSIGETLRGLVNLQ 127  
 QY 124 HLISGQIGRIAGAFDPLESLIEDLDSYNNLRQVPMAGIGAMPALHTLNDHNLIDA 183  
 DB 128 HLIVNNQIGIADAEFEDFLTELDLSYNNLHGLPMDSVRMVNLHQSLDHNLDH 187  
 QY 184 LPPGAFQOLQGLSLDLSNRRLATLAPDPLSRGRDA--EASP--APLVLSFGSNPLHCN 239  
 DB 188 IAGTFADLOGLARLDLSNRLOKLPDPPIFARSQASALTATPPAPPLSFGSNPLHCN 247  
 QY 240 CELLMRLRLARPDDLETCASPGLAGRYFMVAPGEFSCPEPLIARTQRLVMEGQRA 299  
 DB 248 CELLMRLRLERDDLETGSGPGKGRYFMVAREEFVCEPEPLITQHTHKLIVLEGQAT 307  
 QY 300 LRCRALDPAFTMTMVGDBDLVGNSSRARAFPNGTLEIGVTGAGDAGYTCIATNPAGE 359  
 DB 308 LKCAIIGDPSPLIHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCIANAAGE 367  
 QY 360 ATAVERLRLALPHGNSAAGCRGP--SDIASARTAAEGGTLESF-----AVQ 410  
 DB 368 ATANVEVSIQVLPPLSSTSRKAPKSLSDITGSSSKTSRGSGSGGGEPPKSPPERAVL 427  
 QY 411 VTEVATSGIVSWGPGPADVMMFOIQYNSSEDETLIYRIVPASSHFLKHLVPGADY 470  
 DB 428 VSEVTTTSALVKSIVSKAPRVKMYOLQYNSDDEVLIRMIIPASNAFVNNLVSGTGY 487  
 QY 471 DLCLALSPAAGPSDLTATRLLGCAHFTSLPASPLCHALQAHVIGTLTVAAGVVAAL 530

DB 488 DLCLAMWDDTA--TTLTATNIVGCAQPFTRKADYPPQCSMHSQILGTMILVIGIIVATL 546  
 QY 531 LVFTVALIVGRGAGNGRLPLKL-----SHVQSOTNGA---PSTPKAHPRSPP----- 577  
 DB 547 LVFIVILIMVRYK-VCHNEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPGPKVVR 605  
 QY 578 -----PP----- 580  
 DB 606 NELLDFTTASLARSDSSSSSLGSGEAAIGRAPWRIPPSAPRPKPSLDRLMGAFASLDL 665  
 QY 581 -----ORSCSLD 587  
 DB 666 KSQRKEILDSTRTPAGRGACTSARGHHSDBREPLLPAPARASLLPLBCKAKRSHSFD 725  
 QY 588 LGD-----AGCYGARRLGAMARSHSVHGLTG-----AGCRGVGSARLE 631  
 DB 726 MGFPAAMAAAGVVPVGGYSPPRKXSNITWTKRSLSVNGMLPFESDVLVAGRGTRGSSSEWV 785  
 QY 632 ESVY 635  
 DB 786 ESTV 789

Search completed: November 17, 2004, 22:36:45  
 Job time : 77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 22:40:15 ; Search time 141 Seconds  
(without alignments)  
1594.830 Million cell updates/sec

Title: US-10-071-879-10  
Perfect score: 3336  
Sequence: 1 MAPPLLLILLAGAACP...LQAGCRGVGSARLEESVY 635

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1570615 segs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
  - 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
  - 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3319	96.5	778	14	US-10-093-463-192 Sequence 192, App
2	3026.5	90.7	647	14	US-10-093-463-188 Sequence 188, App
3	2965	89.9	565	9	US-09-768-826-58 Sequence 58, Appl
4	2773.5	83.1	566	14	US-10-093-463-190 Sequence 190, App
5	2773	83.1	551	9	US-09-897-214-8 Sequence 8, Appl
6	2475	74.2	468	9	US-09-768-826-40 Sequence 40, Appl
7	2475	74.2	468	11	US-09-823-245-2260 Sequence 2260, Ap
8	1662.5	49.8	832	10	US-09-732-436-24 Sequence 24, Appl
9	1662.5	49.8	832	14	US-10-004-378A-41 Sequence 41, Appl
10	1657.5	49.7	789	14	US-10-004-378A-42 Sequence 42, Appl
11	1655.5	49.6	789	14	US-10-291-172-261 Sequence 261, App
12	1655.5	49.6	789	14	US-10-221-878-261 Sequence 261, App
13	1630	48.9	788	14	US-10-004-378A-43 Sequence 43, Appl

14	1557	46.7	628	9	US-09-815-626-2
15	1557	46.7	628	14	US-10-145-586-61
16	1556	46.6	628	10	US-09-732-436-6
17	1556	46.6	628	14	US-10-221-097-45
18	1546.5	46.4	627	14	US-10-245-752-100
19	1546.5	46.4	627	14	US-10-245-859-100
20	1546.5	46.4	627	14	US-10-245-103-100
21	1546.5	46.4	627	14	US-10-245-107-100
22	1546.5	46.4	627	14	US-10-245-143-100
23	1546.5	46.4	627	14	US-10-245-771-100
24	1546.5	46.4	627	14	US-10-245-851-100
25	1546.5	46.4	627	14	US-10-245-883-100
26	1546.5	46.4	627	14	US-10-237-535-100
27	1546.5	46.4	627	14	US-10-238-183-100
28	1546.5	46.4	627	14	US-10-238-283-100
29	1546.5	46.4	627	14	US-10-238-370-100
30	1546.5	46.4	627	14	US-10-245-147-100
31	1546.5	46.4	627	14	US-10-245-055-100
32	1546.5	46.4	627	14	US-10-245-409-100
33	1546.5	46.4	627	14	US-10-245-730-100
34	1546.5	46.4	627	14	US-10-245-093-100
35	1546.5	46.4	627	14	US-10-246-210-100
36	1546.5	46.4	627	14	US-10-235-196-100
37	1546.5	46.4	627	14	US-10-243-409-100
38	1546.5	46.4	627	14	US-10-243-621-100
39	1546.5	46.4	627	14	US-10-245-880-100
40	1546.5	46.4	627	14	US-10-245-033-100
41	1546.5	46.4	627	14	US-10-243-095-100
42	1546.5	46.4	627	14	US-10-245-185-100
43	1546.5	46.4	627	14	US-10-245-427-100
44	1546.5	46.4	627	14	US-10-245-473-100
45	1546.5	46.4	627	14	US-10-245-770-100

ALIGNMENTS

RESULT 1  
US-10-093-463-192  
Sequence 192, Application US/10093463  
Publication No. US20030208039A1  
GENERAL INFORMATION:  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Shenoy, Suresh  
APPLICANT: Kerkuda, Ramesh  
APPLICANT: Gueev, Vladimir  
APPLICANT: Pochart, Pascal  
APPLICANT: Zhong, Mei  
APPLICANT: Rastelli, Luca  
APPLICANT: Mezes, Peter  
APPLICANT: Smithson, Glenda  
APPLICANT: Guo, Xiaojia  
APPLICANT: Gerlach, Valerie  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Zerhusen, Bryan  
APPLICANT: Tchiernev, Velizar  
APPLICANT: Gangoli, Esna  
APPLICANT: Vernet, Corine  
APPLICANT: Pena, Carol  
APPLICANT: Burgess, Catherine  
APPLICANT: Liu, Xiaohong  
APPLICANT: Spytek, Kimberly  
APPLICANT: Gorman, Linda  
APPLICANT: Spaderna, Steven  
APPLICANT: Voss, Edward  
APPLICANT: Malyankar, Uriel  
APPLICANT: Anderson, David  
APPLICANT: Patturejan, Meeta  
APPLICANT: Miller, Charles  
APPLICANT: Taupier, Raymond J. Jr.  
TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypept

TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.  
FILE REFERENCE: 21402-290A (Cura 590AT)  
CURRENT APPLICATION NUMBER: US/10/093,463  
CURRENT FILING DATE: 2002-06-24  
PRIOR APPLICATION NUMBER: 60/283,675  
PRIOR FILING DATE: 2001-04-14  
PRIOR APPLICATION NUMBER: 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,101  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: 60/287,424  
PRIOR FILING DATE: 2001-04-30  
PRIOR APPLICATION NUMBER: 60/299,027  
PRIOR FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/309,198  
PRIOR FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: 60/281,194  
PRIOR FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: 60/274,194  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,849  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/330,380  
PRIOR FILING DATE: 2001-10-18  
PRIOR APPLICATION NUMBER: 60/275,235  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: 60/288,342  
PRIOR FILING DATE: 2001-05-03  
PRIOR APPLICATION NUMBER: 60/275,578  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 370  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 192  
LENGTH: 778  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-093-463-192

Query Match 96.5%; Score 3219; DB 14; Length 778;  
Best Local Similarity 96.1%; Pred. No. 1.9e-224;  
Matches 617; Conservative 2; Mismatches 5; Indels 18; Gaps 1;

QY 1 MAPPILLILLASGAACPLPCVCONLSISLTLCAHRLGLFVPPNVDRTVEIRLADNFI 60  
DB 1 MAPPILLILLASGAACPLPCVCONLSISLTLCAHRLGLFVPPNVDRTVEIRLADNFI 60  
QY 61 QALGPFRFMTGTVLDTLSNATIRIGARAFGLBLSRLSHLDGNLVELGTSLSLGPV 120  
DB 61 QALGPFRFMTGTVLDTLSNATIRIGARAFGLBLSRLSHLDGNLVELGTSLSLGPV 120  
QY 121 NLGHLISGNOLGRIAPGAPDFLESLEDLDSYNNLRQVWAGIGAMPALHTLNDHL 180  
DB 121 NLGHLISGNOLGRIAPGAPDFLESLEDLDSYNNLRQVWAGIGAMPALHTLNDHL 180  
QY 181 IDALPGAFALQGLSRDLTSLNRIATLAPPLFSRGDAASPAPIVLSFGNPLHCNC 240  
DB 181 IDALPGAFALQGLSRDLTSLNRIATLAPPLFSRGDAASPAPIVLSFGNPLHCNC 240  
QY 241 ELLMLRLARLAPDDIETCASPPGLAGRYFWAVPGEFSGEPPLIARHORLWVLEGGARL 300  
DB 241 ELLMLRLARLAPDDIETCASPPGLAGRYFWAVPGEFSGEPPLIARHORLWVLEGGARL 300  
QY 301 RCRLGDPAPTMWVGPDRLVGNSSRARAPFNGTLEIGVAGDAGGYTCIATNPAGEA 360

DB 301 RCRLGDPAPTMWVGPDRLVGNSSRARAPFNGTLEIGVAGDAGGYTCIATNPAGEA 360  
QY 361 TARVELRVLALPHGNSSAEGGR-----PGSDIASATTAEGEET 402  
DB 361 TARVELRVLALPHGNSSAEGGRPGPRTAPWMEODCRGPPGSDIAASATTAEGEET 420  
QY 403 LSEEPAVQVTEVTATSGVSWGPRPADVMMFOIYVNSSPEDETLIYRIVPASSHFLK 462  
DB 403 LSEEPAVQVTEVTATSGVSWGPRPADVMMFOIYVNSSPEDETLIYRIVPASSHFLK 480  
QY 463 HLPVGDYDLCILALSPAAGPSDLTATRLGCAHFTSLPASPLCHALQAVLGSTLTAV 522  
DB 463 HLPVGDYDLCILALSPAAGPSDLTATRLGCAHFTSLPASPLCHALQAVLGSTLTAV 540  
QY 523 GGTVVAALLVFTVALLVRRGANGRLPLKSHVOSQTMGSGSPTPKAPPPSPRPOR 582  
DB 523 GGTVVAALLVFTVALLVRRGANGRLPLKSHVOSQTMGSGSPTPKAPPPSPRPOR 600  
QY 583 SCSLDLGDAGCYGARRLGAMARRSHSVHGGLGAGCRGVG 624  
DB 601 SCSLDLGDAGCYGARRLGAMARRSHSVHGGLGAGCRGVG 642

RESULT 2  
US-10-093-463-188  
Sequence 188, Application US/10093463  
Publication No. US20030208039A1  
GENERAL INFORMATION:  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Shenoy, Suresh  
APPLICANT: Keskuda, Ramesh  
APPLICANT: Gusev, Vladimir  
APPLICANT: Rochart, Pascal  
APPLICANT: Zhong, Mei  
APPLICANT: Rastelli, Luca  
APPLICANT: Mezes, Peter  
APPLICANT: Smithson, Glenda  
APPLICANT: Guo, Xiaojia  
APPLICANT: Gerlach, Valerie  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Zethusen, Bryan  
APPLICANT: Tcherenev, Velizar  
APPLICANT: Gangolli, Resha  
APPLICANT: Verneet, Corine  
APPLICANT: Pena, Carol  
APPLICANT: Burgess, Catherine  
APPLICANT: Liu, Xiaohong  
APPLICANT: Spytek, Kimberly  
APPLICANT: Gorman, Linda  
APPLICANT: Spaderna, Steven  
APPLICANT: Voss, Edward  
APPLICANT: Malyankar, Uriel  
APPLICANT: Anderson, David  
APPLICANT: Patuturajan, Meera  
APPLICANT: Miller, Charles  
APPLICANT: Taupier, Raymond J. Jr.  
TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypepti  
FILE REFERENCE: 21402-290A (Cura 590AT)  
CURRENT APPLICATION NUMBER: US/10/093,463  
CURRENT FILING DATE: 2002-06-24  
PRIOR APPLICATION NUMBER: 60/283,675  
PRIOR FILING DATE: 2001-04-14  
PRIOR APPLICATION NUMBER: 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,101  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/325,681



```

; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 188
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-093-463-188

```

```

Query Match      90.7%; Score 3026.5; DB 14; Length 647;
Best Local Similarity 93.9%; Pred. No. 1.4e-210;
Matches 589; Conservative 0; Mismatches 21; Indels 17; Gaps 3;

```

```

QY 1 MAPPLILLLASGAACAPLPCVCQNLSESLTLCARGLTFVPPVNDRTVELRLADFI 60
DB 1 MAPPLILLLASGAACAPLPCVCQNLSESLTLCARGLTFVPPVNDRTVELRLADFI 60
QY 61 QALGPDEFRMTGLVDTLTSRNATRTICARAFGDIESTLSLHDGNRLVEIGTSLRGPV 120
DB 61 QALGPDEFRMTGLVDTLTSRNATRTICARAFGDIESTLSLHDGNRLVEIGTSLRGPV 120
QY 121 NLQHLILSGNQLRIAPGAFDDFLESLEDDLSTNNLRQVPWAGIGAMPALHTLNDHNL 180
DB 121 NLQHLILSGNQLRIAPGAFDDFLESLEDDLSTNNLRQVPWAGIGAMPALHTLNDHNL 180
QY 181 IDALPPGAFQOLGQSLRLDITSNRLATLAPDPLFSRGRDAEAPVLVSFGNPLHCNC 240
DB 181 IDALPPGAFQOLGQSLRLDITSNRLATLAPDPLFSRGRDAEAPVLVSFGNPLHCNC 240
QY 241 ELWLRLRLARDDDETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGQATL 300
DB 241 ELWLRLRLARDDDETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGQATL 300
QY 301 RCRLAGDPAPTMHWGVGDDRLVGNSSRARAFPNGLTLEIGTVGADGAGYTCIATNPAGEA 360
DB 301 RCRLAGDPAPTMHWGVGDDRLVGNSSRARAFPNGLTLEIGTVGADGAGYTCIATNPAGEA 360
QY 361 TARVELRLVLAIPHGNSAEGRGPGSDIAASARTAAEGEGTLESSEPAVQVTEVTATSGL 420
DB 361 TARVELRLVLAIPHGNSAEGRGPGSDIAASARTAAEGEGTLESSEPAVQVTEVTATSGL 420
QY 421 VSWGPGPADPVMWFOIQYNSSDEETLIRIVPASSHFLKHLVPGADYDLCLIALSPA 480
DB 421 VSWGPGPADPVMWFOIQYNSSDEETLIRIVPASSHFLKHLVPGADYDLCLIALSPA 480
QY 481 AGPSDLTATRLGCAHFSTLPASPLCHALQAHVGLTIVAVGVVLAALVFTVALVLR 540
DB 481 AGPSDLTATRLGCAHFSTLPASPLCHALQAHVGLTIVAVGVVLAALVFTVALVLR 540

```

```

DB 481 AGPSDLTATRLGCAHFSTLPASPLCHALQAHVGLTIVAVGVVLAALVFTVALVLR 540
QY 541 GRGAGNGRLPLKSHVQSTNGSPSTPKAHPRRSP-----DRPQSCLDYDAG 592
DB 541 GRGAGNGRLPLKSHVQSTNGSPSTPKAHPRRSPBAPAAQLLSGPGRCVRLT----- 596
QY 593 CYGARRLGG-----AMARSSHVHG 614
DB 597 CQAPGRSLQPTTEPLCAKMGAAARGVPEG 623

```

# RESULT 3

```

US-09-768-826-58
; Sequence 58, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PF512P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 58
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (270)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-768-826-58

```

```

Query Match      88.9%; Score 2965; DB 9; Length 565;
Best Local Similarity 99.8%; Pred. No. 3.3e-206;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 71 MTGLVDTLTSRNATRTICARAFGDIESTLSLHDGNRLVEIGTSLRGPVNLQHLILSGN 130
DB 1 MTGLVDTLTSRNATRTICARAFGDIESTLSLHDGNRLVEIGTSLRGPVNLQHLILSGN 130
QY 131 QLGRIACAPDDFLESLEDDLSTNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF 190
DB 131 QLGRIACAPDDFLESLEDDLSTNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF 190
QY 191 QLGQSLRLDITSNRLATLAPDPLFSRGRDAEAPVLVSFGNPLHCNCGLWLRLRLAR 250
DB 191 QLGQSLRLDITSNRLATLAPDPLFSRGRDAEAPVLVSFGNPLHCNCGLWLRLRLAR 250
QY 251 PDLLETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGQATLRCRALGDPAP 310
DB 251 PDLLETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGQATLRCRALGDPAP 310
QY 311 TMHWGVGDDRLVGNSSRARAFPNGLTLEIGTVGADGAGYTCIATNPAGEATARVELRYLA 370
DB 311 TMHWGVGDDRLVGNSSRARAFPNGLTLEIGTVGADGAGYTCIATNPAGEATARVELRYLA 370
QY 371 LPHGNSAEGRGPGSDIAASARTAAEGEGTLESSEPAVQVTEVTATSGVSWGPGPAD 430
DB 371 LPHGNSAEGRGPGSDIAASARTAAEGEGTLESSEPAVQVTEVTATSGVSWGPGPAD 430
QY 431 PVMWFOIQYNSSDEETLIRIVPASSHFLKHLVPGADYDLCLIALSPAAGPSDLTATR 490
DB 431 PVMWFOIQYNSSDEETLIRIVPASSHFLKHLVPGADYDLCLIALSPAAGPSDLTATR 490
QY 491 LIGCAHFSTLPASPLCHALQAHVGLTIVAVGVVLAALVFTVALVLRGRGAGNGRLP 550
DB 491 LIGCAHFSTLPASPLCHALQAHVGLTIVAVGVVLAALVFTVALVLRGRGAGNGRLP 550

```

QY 551 LKLSHVQSGTNGGPEPTPKAHPPPRPPQSGSLDGDAGCTGYARRLGAMARRSHS 610  
DB 481 LKLSHVQSGTNGGPEPTPKAHPPPRPPQSGSLDGDAGCTGYARRLGAMARRSHS 540  
QY 611 VHGGILGAGCGVGGSGARRLEESVY 635  
DB 541 VHGGILGAGCGVGGSGARRLEESVY 565

## RESULT 4

US-10-093-463-190  
; Sequence 190, Application US/10093463  
; Publication No. US20030208039A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Pochart, Pascal  
; APPLICANT: Zhong, Mei  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Mezes, Peter  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Gangoli, Esha  
; APPLICANT: Vermet, Corine  
; APPLICANT: Pena, Carol  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Gormen, Linda  
; APPLICANT: Spaderina, Steven  
; APPLICANT: Vose, Edward  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Anderson, David  
; APPLICANT: Patuturajan, Meera  
; APPLICANT: Miller, Charles  
; APPLICANT: Taupier, Raymond J. Jr.  
; TITLE OF INVENTION: No. US20030208039A1 Anticodons that Bind to Antigenic Polypeptide  
; FILE REFERENCE: 21402-290A (Cura 590AT)  
; CURRENT APPLICATION NUMBER: US/10/093,463  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: 60/283,675  
; PRIOR FILING DATE: 2001-04-14  
; PRIOR APPLICATION NUMBER: 60/338,092  
; PRIOR FILING DATE: 2001-12-03  
; PRIOR APPLICATION NUMBER: 60/274,281  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/274,101  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/325,681  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/304,354  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/279,995  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/294,899  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/287,424  
; PRIOR FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 60/299,027  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/309,198  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/281,194  
; PRIOR FILING DATE: 2001-04-04

; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/274,849  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/330,380  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/275,235  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/288,342  
; PRIOR FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 60/275,578  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 370  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 190  
; LENGTH: 566  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-093-463-190

Query Match 83.1%; Score 2773.5; DB 14; Length 566;  
Best Local Similarity 95.4%; Pred No. 2.5e-192;  
Matches 537; Conservative 4; Mismatches 11; Indels 11; Gaps 2;

QY 1 MAPPLILLASGAACPLPCVCONLSLSLTLCAHRIILFVPPVNDRTVEIRLADNFI 60  
DB 1 MAPPLILLASGAACPLPCVCONLSLSLTLCAHRIILFVPPVNDRTVEIRLADNFI 60  
QY 61 QALGPPDFRMTGLVDLTISRNAITRIGARAGDLESLSLHLDGRLVEIGTSGIRGPV 120  
DB 61 QALGPPDFRMTGLVDLTISRNAITRIGARAGDLESLSLHLDGRLVEIGTSGIRGPV 120  
QY 121 NIQHILISGNOLGRIAPGAFDPLESLRDLDSYNNLRQVPWAGIGAMPALHTLNDHNL 180  
DB 121 NIQHILISGNOLGRIAPGAFDPLESLRDLDSYNNLRQVPWAGIGAMPALHTLNDHNL 180  
QY 181 IDALPPGAFQGLSRLDLSNRRLATLAPDPLFSKGRDAESPAPLVLSFGNPLHCNC 240  
DB 181 IDALPPGAFQGLSRLDLSNRRLATLAPDPLFSKGRDAESPAPLVLSFGNPLHCNC 240  
QY 241 ELLMLRRLARPDDLETGASPPGLAGRYFPAVVBGEESCEPPLIARHTORLWVEGGRATL 300  
DB 241 ELLMLRRLARPDDLETGASPPGLAGRYFPAVVBGEESCEPPLIARHTORLWVEGGRATL 300  
QY 301 RCRAIGDPAPTMHWGPPDDRLVGNSSRARAFPGTLEIVTGAGDAGYTCIATNPAGEA 360  
DB 301 RCRAIGDPAPTMHWGPPDDRLVGNSSRARAFPGTLEIVTGAGDAGYTCIATNPAGEA 360  
QY 361 TARVELRVLALPHGNSAEGGPPGSDIAASARTAAEGGTLESEPAVQVTEVTATSGL 420  
DB 361 TARVELRVLALPHGNSAEGGPPGSDIAASARTAAEGGTLESEPAVQVTEVTATSGL 420  
QY 421 VSWGGRPADPVMFQIYNSSSEDTLIRIVASSHHFLIKLIVGADVDLCILALSPA 480  
DB 421 VSWGGRPADPVMFQIYNSSSEDTLIRIVASSHHFLIKLIVGADVDLCILALSPA 480  
QY 481 AGPSDLTATRLIGCAFSTLPASPLCHALOAHVIGTLTVAVGVVYAAALVFTVALVR 540  
DB 481 AGPSDLTATRLIGCAFSTLPASPLCHALOAHVIGTLTVAVGVVYAAALVFTVALVR 540  
QY 541 GRG-----AGNGRLPLKL 553  
DB 540 WRGLPLPWSGGQNRKGSILPLQV 562

RESULT 5  
US-09-897-214-8  
; Sequence 8, Application US/09897214  
; Patent No. US20020076779A1  
; GENERAL INFORMATION:  
; APPLICANT: Thayer, Edward C.  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Presnell Scott R.

; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,  
 ; FILE REFERENCE: Zlrr8, and Zlrr9  
 ; CURRENT APPLICATION NUMBER: US/09/897,214  
 ; PRIOR FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: 60/215,446  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 551  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-897-214-8

Query Match 83.1%; Score 2773; DB 9; Length 551;  
 Best Local Similarity 100.0%; Pred. No.2,6e-192;  
 Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPPILLILLASGAACPLPCVCONLSESLSTLCARGLLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCONLSESLSTLCARGLLFVPPNVDRTVELRLADNFI 60  
 QY 61 QALGPDEFNMTGLVLTLSRNAITRIQARAFGDLISLGLHLDGNRLVELGTSLRGV 120  
 DB 61 QALGPDEFNMTGLVLTLSRNAITRIQARAFGDLISLGLHLDGNRLVELGTSLRGV 120  
 QY 121 NLOHLISGNOLGRIAPGAFDDLESLDLSYNNLROYPMAGIGAMPALHTLNDHL 180  
 DB 121 NLOHLISGNOLGRIAPGAFDDLESLDLSYNNLROYPMAGIGAMPALHTLNDHL 180  
 QY 181 IDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 DB 181 IDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 QY 241 ELLMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCBPLIARHTOCLMWLEGGATL 300  
 DB 241 ELLMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCBPLIARHTOCLMWLEGGATL 300  
 QY 301 RCRALGDPAPTMHWVGPDDRLVGNSSRAAFPNGTLEIGVTSAGDAGGYTCTATNPAGEA 360  
 DB 301 RCRALGDPAPTMHWVGPDDRLVGNSSRAAFPNGTLEIGVTSAGDAGGYTCTATNPAGEA 360  
 QY 361 TARVELRLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESPPAVQTEVTATSGL 420  
 DB 361 TARVELRLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESPPAVQTEVTATSGL 420  
 QY 421 VSMGGRPADPYWMQIQYNSSDETLIRIVPASSHHFLKHLVPGADYDCLIALSPA 480  
 DB 421 VSMGGRPADPYWMQIQYNSSDETLIRIVPASSHHFLKHLVPGADYDCLIALSPA 480  
 QY 481 AGPSDLTATRLIGCAHFTLPPASPLCHALQAHVLGGTLTVAAGVVAL 530  
 DB 481 AGPSDLTATRLIGCAHFTLPPASPLCHALQAHVLGGTLTVAAGVVAL 530

RESULT 6  
 ; US-09-768-826-40  
 ; Sequence 40, Application US/09768826  
 ; Patent No. US20020012966A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shi et al.  
 ; TITLE OF INVENTION: 18 human secreted proteins  
 ; FILE REFERENCE: PF512P1  
 ; CURRENT APPLICATION NUMBER: US/09/768,826  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: PCT/US00/22350  
 ; PRIOR FILING DATE: 2000-08-15  
 ; PRIOR APPLICATION NUMBER: 60/148,759  
 ; NUMBER OF SEQ ID NOS: 61  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 40

; LENGTH: 468  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-768-826-40

Query Match 74.2%; Score 2475; DB 9; Length 468;  
 Best Local Similarity 99.8%; Pred. No.8,3e-171;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227  
 DB 1 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227  
 QY 228 VLSFGSNPLHCNCELMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCBPLIARHT 287  
 DB 61 VLSFGSNPLHCNCELMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCBPLIARHT 287  
 QY 288 QRLMWLEGGATLRCRAGDPAPTMHWVGPDDRLVGNSSRAAFPNGTLEIGVTSAGDAG 347  
 DB 121 QRLMWLEGGATLRCRAGDPAPTMHWVGPDDRLVGNSSRAAFPNGTLEIGVTSAGDAG 347  
 QY 348 GTTCTATNPAGEATARVELRLVIALPHGNSAEGGRPGPSDIAASARTAAEGGTLESSE 407  
 DB 181 GTTCTATNPAGEATARVELRLVIALPHGNSAEGGRPGPSDIAASARTAAEGGTLESSE 407  
 QY 408 AVQVTEVTATSGLVSMGGRPADPYWMQIQYNSSDETLIRIVPASSHHFLKHLVPG 467  
 DB 241 AVQVTEVTATSGLVSMGGRPADPYWMQIQYNSSDETLIRIVPASSHHFLKHLVPG 467  
 QY 468 ADVDCIALLSPPAGPSDLTATRLIGCAHFTLPPASPLCHALQAHVLGGTLTVAAGVVAL 527  
 DB 301 ADVDCIALLSPPAGPSDLTATRLIGCAHFTLPPASPLCHALQAHVLGGTLTVAAGVVAL 527  
 QY 528 AALLVFTVALIVRGAGNGRLPLKLSHVQSQTNCGSPSTPAHPPRSPPPRQSCSLD 587  
 DB 361 AALLVFTVALIVRGAGNGRLPLKLSHVQSQTNCGSPSTPAHPPRSPPPRQSCSLD 587  
 QY 588 LGDACCYGARPLGGAMARRSHSVHGLLAGCCRGVGSAAERLESYV 635  
 DB 421 LGDACCYGARPLGGAMARRSHSVHGLLAGCCRGVGSAAERLESYV 635

RESULT 7  
 ; US-09-833-245-2260  
 ; Sequence 2260, Application US/09833245  
 ; Publication No. US20040010134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546PCT  
 ; CURRENT APPLICATION NUMBER: US/09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2260  
 ; LENGTH: 468  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-833-245-2260

Query Match 74.2%; Score 2475; DB 11; Length 468;  
 Best Local Similarity 99.8%; Pred. No.8,3e-171;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 168 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227  
 DB 1 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227



TITLE OF INVENTION: Methods of Using the Same  
FILE REFERENCE: 21402-179  
CURRENT APPLICATION NUMBER: US/10/004,378A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 60/242,882  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/242,765  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/300,206  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/242,789  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/242,768  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/242,767  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/243,622  
PRIOR FILING DATE: 2000-10-26  
PRIOR APPLICATION NUMBER: 60/273,047  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: 60/243,591  
PRIOR FILING DATE: 2000-10-26  
PRIOR APPLICATION NUMBER: 60/243,950  
PRIOR FILING DATE: 2000-10-27  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 41  
LENGTH: 832  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-004-378A-41

Query Match 49.8%; Score 1662.5; DB 14; Length 832;  
Best Local Similarity 46.4%; Pred. No. 1,3e-111;  
Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

9 LIAAGAA-----ACPLPCYQOMLSLSLSTLCARGLFVPPVNDRTVELRLADNFIAL 63  
Db 51 LIAFGAFVAVNDACKRYCVCQMLSESLGTLCPKGLFVPPIDRRVELRLGNGFIHI 110  
Qy 64 GPPDPRNTGIVDVLTSNRNATIRIGARAFGDLSESLHLDONRIVEIGTGLRPPVNIQ 123  
Db 111 SRQDPRANTGIVDVLTSNRNATIRIGARAFGDLSESLHLDONRIVEIGTGLRPPVNIQ 170  
Qy 124 HLISGNQGRIPAGAFDPLESLDLISYNNLRQVPAAGAMPALHTTLDHNLDA 183  
Db 171 HLIVNNGQGGIADAEFDFLTLEDLDLSTYNNHGLFWDSVRRVNIHQSLDENLDDH 230  
Qy 184 LPPGAFQGLSRDLDTSNRLATITAPPPLFSRGRDA--EASP-APVLSPSGNPLHCN 239  
Db 231 IAEGLFADLOKLARLDLTSNRLOKLPDPPIFARSGASALATPPAPPLSPFSFGNPLHCN 290  
Qy 240 CELMLRLRLAPDDLETGASPPGLAGRFMAVPEGESECEPPLIARHQLRWLEFGQAT 299  
Db 291 CELMLRLRLDDDDLETGASPPGLAGRFMAVPEGESECEPPLIARHQLRWLEFGQAT 350  
Qy 300 LRCRALGDPAPTMVWGPDDRLVGNSSRRARAFPNGTLEIGVYAGDAGYCIATNPGE 359  
Db 351 LKCAIGDPSPLIHVVAPDDRLVGNSSRRARAFPNGTLEIGVYAGDAGYCIATNPGE 410  
Qy 360 ATAVELRVALLPHGSSNABGRRGP--SDIAMSARTAAEGGTLESSEP-----AVQ 410  
Db 411 ATAVELRVALLPHGSSNABGRRGP--SDIAMSARTAAEGGTLESSEP-----AVQ 470  
Qy 411 VTEWTATGCLVSWGGRPADPVMFMQIOYNSSEDETLIRVPPASHHFLKHLVPGADY 470  
Db 471 VSEVTTTSALVKMSVSPAPVKKYQIQIQYNSDEDEVLIRYMI PASNKAFFVNNIIVSGGY 530  
Qy 471 DLCLLAISPAAGPDLTATRLILGCAHFSTLPASPLCEALQAVLGTLTVAGVLAAL 530  
Db 531 DLCLLAISPAAGPDLTATRLILGCAHFSTLPASPLCEALQAVLGTLTVAGVLAAL 589

Qy 531 IVEFTVALYVRGAGNGRPLKLU-----SHVQSOTNG-----PSFTPKAPRSP----- 577  
Db 590 IVEFTVALYVRGAGNGRPLKLU-----SHVQSOTNG-----PSFTPKAPRSP----- 648  
Qy 578 -----PRP----- 580  
Db 649 NELDEFTASLARASDSSSSSSSGSGBAAGLGRAPWRIPPSAPRPSLDRIMGAFASLDL 708  
Qy 581 -----ORCSLD 587  
Db 709 KQQRKEELDSRTPAAGGAGTSARGHSDREBPLGPPAARASLLPLPLEGKAKRSHSFD 768  
Qy 588 LQD-----AGCYGARRLGAMARSHSVHGLG-----AGCRGVGSAERLE 631  
Db 769 MGDFPAAAGAGVPPGYSPPRKYVSNITKRSLSVNGMLPFRESDLVGARGTGGSSEWM 828  
Qy 632 ESVV 635  
Db 829 ESTV 832

## RESULT 10

US-10-004-378A-42  
Sequence 42, Application US/10004378A  
Publication No. US20030228301A1  
GENERAL INFORMATION:

APPLICANT: Li, Li  
APPLICANT: Futak, Kazaryna  
APPLICANT: Perna, Amanda  
APPLICANT: Patuxajan, Meera  
APPLICANT: Shinkets, Richard A  
APPLICANT: Guo, Xiaojia Sasha  
APPLICANT: Casman, Stacie J  
APPLICANT: Burgess, Catherine E  
APPLICANT: Malyankar, Uriel M  
APPLICANT: Tchernev, Velizar T  
APPLICANT: Vernet, Corinne A  
APPLICANT: Spytek, Kimberly A  
APPLICANT: Agge, Michele  
APPLICANT: Rascelli, Luca  
APPLICANT: Shenoy, Suresh G  
APPLICANT: Grosche, William M  
APPLICANT: Alschbrook II, John P  
APPLICANT: Lepley, Denise M  
APPLICANT: Edinger, Valerie  
APPLICANT: MacDougall, John R  
APPLICANT: Gunther, Erik  
APPLICANT: Peyman, John A  
APPLICANT: Stone, David J  
APPLICANT: Ellerman, Karen  
APPLICANT: Gangolli, Esra A  
TITLE OF INVENTION: No. US20030228301A1 Human Proteins, Polynucleotides Encoding T  
FILE REFERENCE: 21402-179  
CURRENT APPLICATION NUMBER: US/10/004,378A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 60/242,882  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/242,765  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/300,206  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/242,789  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/242,768  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/242,767  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 60/243,622  
PRIOR FILING DATE: 2000-10-26  
PRIOR APPLICATION NUMBER: 60/273,047  
PRIOR FILING DATE: 2001-03-02

```

; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 789
; TYPE: PR
; ORGANISM: Cynomolgus monkey
US-10-004-378A-42

Query Match      49.7%; Score 1657.5; DB 14; Length 789;
Best Local Similarity 46.3%; Pred. No. 2.8e-111;
Matches 363; Conservative 82; Mismatches 180; Indels 159; Gaps 13;

QY 9 LIAAGAA-----ACPLPCVCONSESLSTLCARGLLFPVPVDRRTVELRLADNFIQAL 63
DB 8 LIAFGMAFAVVDACPKYCVCQONSESLGTLCPKGLFPFPDIDRTVELRLGNGFIHI 67
QY 64 GPPDFRNMGTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVELGTSLRGPNVLIQ 123
DB 68 SRQDFANMTGLVDLTLSRNTISHIQPFSLDELSLSLHDSRLSLGDLTGLVNLQ 127
QY 124 HLISGNQIGRIAPGAFDDFLESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNLIDA 183
DB 128 HLIVNNNOIGLADAEFEEDFLTLTEDLDSYNNLHGLPMDSVARMVNLHQLSDHNLIDH 187
QY 184 LPPGAFAOIGQSLRLDLSRNAITLAPDPLFSRGRDA--EASP--APVLVSFGNPLHCN 239
DB 188 IAEGTADLOKRLDLTSLRLOKLPDPPIFARQASALTATPAPPLSFGNPLHCN 247
QY 240 CELLMRLRLARPDDLETCASPPGLAGRYFWAVEGEFSCPEPLIARHTORLWVLEQORAT 299
DB 248 CELLMRLRLERDDLETCSPGGLKGRYFWAVEEVEFCEPLITQHTHLVLLEGQAT 307
QY 300 LRCRALGDPAPTMWVGPDDRILVGNSSRAAFPNGLTLEICVTGAGAGYTCIATNPAGE 359
DB 308 LKCKALGDPBPLIHWVAPDDRLVGNSSRTAVYNGTIDITTSQDSGATCIAANAGE 367
QY 360 ATARVELRYVALPHGNGSSAEGRPGP--SDIAASARTAAEGTLESEP-----AVQ 410
DB 368 ATTVEXSTIQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGSGGSGEPKSPPERAVL 427
QY 411 VTEVATSGLVSWGPGRPADPVMMFOIQYNSSEDETLIYRIVPASHHFLKHLVPGADY 470
DB 428 VSEVTTTSLAKWVSXKSTPRVVMYQIQYNSCDEVL IYRMI PASNKA FVYNNLVSGTGY 487
QY 471 DLCLLALSPAAGSDLTATRLGCAHFSTLPASPLCHALAHVUGTLTVAAGVLYAAL 530
DB 488 DLCLVLAAMWDDTA--TTLTATNIVGCAQFTRADYPOCCSMHSQILGTMILVIGIIVATL 546
QY 531 LVFTVALLVGRGAGNRLPLKL-----SHVOSQTNCG-----PSPTPKAHPRSP 577
DB 547 LVFTVILMVRYK--VCNHEAPSKMAAAVSNVYSQTNQAPPPSSAPRGAPPGPKVYVR 605
QY 578 -----PRP----- 580
DB 606 NELLDFTASLAPASDSSSSSSISGGAAGLGRAPWRLPPSAPRPKPSLDRIMGAFASLDL 665
QY 581 -----ORSCSLD 587
DB 666 KSKORKEHLDSRTPPAGGAGTSARGHSDREPLIGPPAARASLLPLPLGKAKKSSSFD 725
QY 588 LGD-----AGCYGARRLGAMARRSHVHGGLG-----AGCRGVGGSABRL 631
DB 726 MGDFAAALAAAGGVVPGISPRRVSNTYTKSLSTVNGMLPREESDVGAGGTFGSSWMV 785
QY 632 ESVV 635
DB 786 ESTV 789

```

```

RESULT 11
US-10-291-172-261
; Sequence 261, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 261
; LENGTH: 789
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-291-172-261

```

```

Query Match      49.6%; Score 1655.5; DB 14; Length 789;
Best Local Similarity 46.3%; Pred. No. 3.9e-111;
Matches 363; Conservative 83; Mismatches 179; Indels 159; Gaps 13;

```

```

QY 9 LIAAGAA-----ACPLPCVCONSESLSTLCARGLLFPVPVDRRTVELRLADNFIQAL 63
DB 8 LIAFGMAFAVVDACPKYCVCQONSESLGTLCPKGLFPFPDIDRTVELRLGNGFIHI 67
QY 64 GPPDFRNMGTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVELGTSLRGPNVLIQ 123
DB 68 SRQDFANMTGLVDLTLSRNTISHIQPFSLDELSLSLHDSRLSLGDLTGLVNLQ 127
QY 124 HLISGNQIGRIAPGAFDDFLESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNLIDA 183
DB 128 HLIVNNNOIGLADAEFEEDFLTLTEDLDSYNNLHGLPMDSVARMVNLHQLSDHNLIDH 187
QY 184 LPPGAFAOIGQSLRLDLSRNAITLAPDPLFSRGRDA--EASP--APVLVSFGNPLHCN 239
DB 188 IAEGTADLOKRLDLTSLRLOKLPDPPIFARQASALTATPAPPLSFGNPLHCN 247
QY 240 CELLMRLRLARPDDLETCASPPGLAGRYFWAVEGEFSCPEPLIARHTORLWVLEQORAT 299
DB 248 CELLMRLRLERDDLETCSPGGLKGRYFWAVEEVEFCEPLITQHTHLVLLEGQAT 307
QY 300 LRCRALGDPAPTMWVGPDDRILVGNSSRAAFPNGLTLEICVTGAGAGYTCIATNPAGE 359
DB 308 LKCKALGDPBPLIHWVAPDDRLVGNSSRTAVYNGTIDITTSQDSGATCIAANAGE 367
QY 360 ATARVELRYVALPHGNGSSAEGRPGP--SDIAASARTAAEGTLESEP-----AVQ 410
DB 368 ATTVEXSTIQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGSGGSGEPKSPPERAVL 427
QY 411 VTEVATSGLVSWGPGRPADPVMMFOIQYNSSEDETLIYRIVPASHHFLKHLVPGADY 470
DB 428 VSEVTTTSLAKWVSXKSTPRVVMYQIQYNSCDEVL IYRMI PASNKA FVYNNLVSGTGY 487
QY 471 DLCLLALSPAAGSDLTATRLGCAHFSTLPASPLCHALAHVUGTLTVAAGVLYAAL 530
DB 488 DLCLVLAAMWDDTA--TTLTATNIVGCAQFTRADYPOCCSMHSQILGTMILVIGIIVATL 546
QY 531 LVFTVALLVGRGAGNRLPLKL-----SHVOSQTNCG-----PSPTPKAHPRSP 577
DB 547 LVFTVILMVRYK--VCNHEAPSKMAAAVSNVYSQTNQAPPPSSAPRGAPPGPKVYVR 605

```



QY 578 -----PRP----- 580  
Db 606 NEILDFTASLARSDDSSSSSLGSGEAGLGRAPWRIIPSPAPRPKPSLDRLMGAFASLDL 665  
QY 581 -----ORCSLD 587  
Db 666 KSQRKEELDSRTPAAGAGTSARGHHSREPLLGPPAARASLLPLLEGAKRSHSFD 725  
QY 588 LGD-----ACGYARRLGAMARRSHSVHGLG-----ACGRVGSASARLE 631  
Db 726 MGDFAAAAAGGVPGGYSPPRKYSNIWTKRSLSVNGMLLPFEESDLVGARGTGSSEWVW 785  
QY 632 ESTV 635  
Db 786 ESTV 789

## RESULT 12

US-10-221-278-261  
; Sequence 261, Application US/10221278  
; Publication No. US20040034208A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-045  
; CURRENT APPLICATION NUMBER: US/10/221,278  
; PRIOR FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: 09/693,267  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/665,363  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: 09/616,847  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 09/596,193  
; PRIOR FILING DATE: 2000-06-17  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; NUMBER OF SEQ ID NOS: 752  
; SEQ ID NO 261  
; LENGTH: 789  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-221-278-261

Query Match 49.6%; Score 1655.5; DB 15; Length 789;  
Best Local Similarity 46.3%; Pred No. 3.9e-111;

Matches 363; Conservative 83; Mismatches 179; Indels 159; Gaps 13;

QY 9 LIASGAA-----ACPLPCVQNTLSSTLCARGLTFVPPVNDRTVELRLADFIQAL 63  
Db 8 LIAFGMAFVAVDACPKYCVQNLSSSLGTLCPSSKGLVPPDIDRITVELLGNFIHI 67  
QY 64 GPPFRMNTGLVDLTLSNATRTGARAFCGLSIRSHLDGNRLVELGSLRGPVNLQ 123  
Db 68 SRDPANNTGLVDLTLSNATRTSHIQPSFLDESIRSHLDGNRLVELGSLRGPVNLQ 127  
QY 124 HLISGNQIGRIAPGAFDPLESLEDLSTYNNIROYWAGIGAMPALHTLNDHNLIDA 183  
Db 128 HLIVNNOGLGIDAEAFEDLITLEDLSTYNNIHLGFWDSVRKVNHLQSLDHNLDH 187  
QY 184 LPPGAFQGLSLRLDLSNRLATLADPLFSRGRDA--EASP--APLVLSFGSNPLHCN 239  
Db 188 IAEGTFAQLKTLARLDLSNRLQLPDPPIFARQASALTLTPAPRLSFEFGNPLHCN 247  
QY 240 CELIMRLIARPDLETCASPGIAGRYTAVPGEFSCEPPLIARHTQRLVLEGGQAT 299  
Db 248 CELIMRLIARPDLETCASPGIAGRYTAVPGEFSCEPPLIARHTQRLVLEGGQAT 307  
QY 300 LRCRALDPAFTMHWGPDRLVGNSSRARAFPNGTLEIGVTAGDAGGYTCIATNPAGE 359

Db 308 LKCAIGDPSPLIHVWAPDDRLVGNSSRTAVYNDNGTPIFITTSQDSGAFCTIAANAGE 367  
QY 360 ATARVRLRLIALPHGNSNAEGGRPCP--SDIAASRLTAEGGTLESEP-----AVQ 410  
Db 368 ATAMVEVSIVOLPHLNSNSTSRTPAPKSRSLDLTGSSKTSRGSGGCGEPPKSPPERAVL 427  
QY 411 VREVTATSGVSGPGRPADPVMFOIYVNSSSEDELTIRIYVASSHHPLKHLVPGADY 470  
Db 428 VSEVITTSALVKNVSKSPRRKMTQLYNCSDEVLITRMIIPASKAFVNNLVSGTGY 487  
QY 471 DLCLLALSPAAGPSDLTATRLGCAHFTLPASPLCALQAHVIGTLTVAVGVVYAL 530  
Db 488 DLCLAMWDDTA--TLTATNIVGCAQFTFKAYIPQCGSHSQTIGSTMILVIGITVATL 546  
QY 531 LVFTVALLVGRGAGNGRLPLKL-----SHVQQTNGC-----PSPPKAPPPSP 577  
Db 547 LVFTVILWVYK--VCNHEAPSKMAAVALSVQTMGACPPPPSAPAGAPPGQPPKVVVR 605  
QY 578 -----PRP----- 580  
Db 606 NEILDFTASLARSDDSSSSSLGSGEAGLGRAPWRIIPSPAPRPKPSLDRLMGAFASLDL 665  
QY 581 -----ORCSLD 587  
Db 666 KSQRKEELDSRTPAAGAGTSARGHHSREPLLGPPAARASLLPLLEGAKRSHSFD 725  
QY 588 LGD-----ACGYARRLGAMARRSHSVHGLG-----ACGRVGSASARLE 631  
Db 726 MGDFAAAAAGGVPGGYSPPRKYSNIWTKRSLSVNGMLLPFEESDLVGARGTGSSEWVW 785  
QY 632 ESTV 635  
Db 786 ESTV 789

## RESULT 13

US-10-004-378A-43  
; Sequence 43, Application US/10004378A  
; Publication No. US20030228301A1  
; GENERAL INFORMATION:

; APPLICANT: Li, Li  
; APPLICANT: Furtak, Kazazyna  
; APPLICANT: Perrina, Amanda  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Guo, Xiaojia Sasma  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Malyankar, Urfiel M  
; APPLICANT: Tchervet, Vellizar T  
; APPLICANT: Vernet, Corinne A  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Agee, Michele  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Grose, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Edinger, Scholmit  
; APPLICANT: MacDougall, John R  
; APPLICANT: Peyman, John A  
; APPLICANT: Gunther, Erik  
; APPLICANT: Stone, David J  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gangoli, Esha A  
; TITLE OF INVENTION: No. US20030228301A1el Human Proteins, Polynucleotides Encoding T  
; FILE REFERENCE: 21402-179  
; CURRENT APPLICATION NUMBER: US/10/004,378A  
; PRIOR FILING DATE: 2001-10-24  
; PRIOR APPLICATION NUMBER: 60/242,882

PRIOR APPLICATION NUMBER: 60/242,765  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/300,206  
 PRIOR FILING DATE: 2001-06-22  
 PRIOR APPLICATION NUMBER: 60/242,789  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/242,768  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/242,767  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/243,622  
 PRIOR FILING DATE: 2000-10-26  
 PRIOR APPLICATION NUMBER: 60/273,047  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: 60/243,591  
 PRIOR FILING DATE: 2000-10-26  
 PRIOR APPLICATION NUMBER: 60/243,950  
 PRIOR FILING DATE: 2000-10-27  
 Remaining Prior Application data removed - See file Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 191  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 43  
 LENGTH: 788  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-10-004-378A-43

Query Match 48.9%; Score 1630; DB 14; Length 788;  
 Best Local Similarity 46.0%; Pred. No. 2,7e-109;  
 Matches 360; Conservative 84; Mismatches 181; Indels 158; Gaps 13;

QY 9 LLAGAA-----ACPLPCVQONISESLSTLCARGLLFVPPNVDRTVELRLADNFIOAL 63  
 DB 8 LIAFGAFAVVDACPYXCVCQONISESLSTLCPSKRLLFVPPDIDRTVELRLGNGFIHH 67  
 QY 64 GPPDPNNMGLVDTLSRNAITRIGARAFGDLSTLSLHDGRVLELGTSLRGPNVLO 123  
 DB 68 GRODPNMMGLVDTLSRNAITRIGARAFGDLSTLSLHDGRVLELGTSLRGPNVLO 127  
 QY 124 HLISGNOIGRIAPGAFDDFLESLDLDLSYNNLRQVPMAGIGAMPALHTLNDHNLDA 183  
 DB 128 HLIVNNQIGRIDDAFEDFLTLTLEDLDSYNNLHGLPMDSVRRMVLHQLSDHNLDDH 187  
 QY 184 LPPGAPOQLSRLDITSNRLATLAPDLFSRGRDA--EASP--APLVLSFGNPLHCN 239  
 DB 188 IAGTFADLOKLRLDITSNRLQKLPDPPIFARSQASLTATPAPPLSFGNPLHCN 247  
 QY 240 CELLMLRRLARPDDLETCASPPGLAGRYFWAVPEGEFSCPEPLIARTTORLMTLEGORAT 299  
 DB 248 CELLMLRRLERDDLETCSPGLSFKRYFWHIREEFVCEPPLITQHTHKLVLBEGAAT 307  
 QY 300 LRCRALDPAPTMHWGPDRLVGNSSRAAPFNGTLEIVTGAAGAGYTCIATNPAGE 359  
 DB 308 LKCAIGDPSPLIHWVAPDRLVGNSSRTAVYDNGTLDITTSQDSGPFCTIAANNAAGE 367  
 QY 360 ATARVRLVALPFGGSSAEGGRPG--SDIAASARTABEGTLESEP-----AVQ 410  
 DB 368 ATATVESIVQLPPLISNSTSRMAPPKSRLDITGSSSRGGSAGGEPKSTPEPAVL 427  
 QY 411 VTEVTATSGLVSMGPRPADPVMWFQIQYNSSEDETLIYRIVPASHHFLKHLVPGADY 470  
 DB 428 VSDVTTTSLVKSASAPRYVMYQIQYNSDDEVLITMIPASNAFVYNNLVSTGY 487  
 QY 471 DLCLALSPAAGSDTLATRLIGCAHFSTLPASPLCHALQAHVIGTLTVAAGVLYAAL 530  
 DB 488 DLCLVAMWDITA--TLTATNIVGCAOFETKADYPOCOSMHSQIKGTMIIIVTIGIITATL 546  
 QY 531 LVFVALLVGRGNGRLPLKL-----SHVQSQTNGGSP-----TPKXHP-- 573  
 DB 547 LVFVILIMVRYK--VCNHDTPGKMAAATVSNVYQSQTNSQBPPLGIVGQLPQAPKVV 605  
 QY 574 -----RSPPPR----- 580

DB 606 RNEIMDFSTSLARACDSSSSSLSGSEAAHGLRGWPRLPPAPRPSPSLDRLMGAFASID 665  
 QY 581 -----ORSCSL 586  
 DB 666 LKSORKEELDSRTPAGAGCTSSRGHSHDREPLIGPATRASLLPLLEBKAKRSHSF 725  
 QY 587 DLGD-----AGCYGARRLGAMARRSHSHVHGLIG-----ACRGVGSAAERLEE 632  
 DB 726 DWGDFAAAAAIVPGGYSPPRRVSNITWTRSLSYNNMLLPFEESDLVGARGTGSSEWYWE 785  
 QY 633 SYV 635  
 DB 786 STV 788

RESULT 14  
 US-09-815-626-2  
 Sequence 2, Application US/09815626  
 Patent No. US2002076752A1  
 GENERAL INFORMATION:  
 APPLICANT: Gluckmann, Maria Alexandra  
 TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT  
 TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF  
 FILE REFERENCE: 10448-031001  
 CURRENT APPLICATION NUMBER: US/09/815,626  
 CURRENT FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: US 60/191,863  
 PRIOR FILING DATE: 2000-03-24  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 628  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-815-626-2

Query Match 46.7%; Score 1557; DB 9; Length 628;  
 Best Local Similarity 54.0%; Pred. No. 4e-104;  
 Matches 329; Conservative 71; Mismatches 175; Indels 34; Gaps 8;

QY 4 PLILILLAGAAA-----CPPLPCVQONISESLSTLCARGLLFVPPNVDRTVEL 53  
 DB 5 PLILILPLAPASSPPOSATPSPCPRRCOTQSLPSVLCPGAALLFVPPSLDRPAEL 64  
 QY 54 RLADNFIOALGPPDPNNMGLVDTLSRNAITRIGARAFGDLSTLSLHDGRVLELGT 113  
 DB 65 RLADNFIASRRDLANMTGLHLSTSRITRIVAAAGAFADLALALHLDGRVLSIGE 124  
 QY 114 GSLRGVNTLOHLISGNOIGRIAPGAFDDFLESLDLDLSYNNLRQVPMAGIGAMPALHT 173  
 DB 125 GQLRGVNLRLHILISNNQLAALAAAGALDDCAETLEDLDSYNNLEQLPWEALGRIGVNT 184  
 QY 174 LINDHNLIDALPPGAPOQLSRLDITSNRLATLAPDLFSR-----GRPAASPAVLV 229  
 DB 185 LGIDHNLIASVPPGARSRLKRLALDMTSNRLTTIPDPLFSRLPLARRGSPASALV 244  
 QY 230 SFGNPLHNCCELLMLRRLARPDDLETCASPPGLAGRYFWAVPEGEFSCPEPLIARTTOR 289  
 DB 245 AFGNPLHNCCELLMLRRLAREDDLETCASPPALGGYFWAVGEFEFVCEPVPVTHRSP 304  
 QY 290 LMTLEGORATLRCRALDPAPTMHWGPDRLVGNSSRAAPFNGTLEIVTGAAGAGY 349  
 DB 305 LAVPAGRPALRCRAVDDPEPRVWVSPQGRLLGNSSRAAPFNGTLEIVTGEDGIGIF 364  
 QY 350 TCIATNPAGATARVELRLALPH--CGNSSAEGGRPG-----PSDIAASARTAAEG 399  
 DB 365 TCIANNAAGATAVELTVPPPEPQLANSTSCDPPDGPDPALTTPSAASAKAVADTG 424  
 QY 400 EGTLESEPAVOYEVATSGLVSMGPRPADPVMWFQIQYNSSEDETLIYRIVPASSHHF 459  
 DB 425 PPT---DRGQVTEHGTALVQMPDRPIPIGRIMYQIQYNSADDITLVYRMTFAESRSF 481  
 QY 460 LKHLVPGADVDLCLIALSPAAGSDTLATRLIGCAHFSTLPASPLCHALQAHVIGTLT 519

Db 482 LITDLSGRTYDLCVLAIVEGSA-TGLTATRPVGARFSTEPALRPGAHAFLLGGTMI 540  
QY 520 VAVGVIVAAALVFTVALVR-----GRGANGRLPLKLSHVOSQTNCGSPPTPKAHPR 574  
Db 541 IALGVIVASVIVFTVFLMRKYKHGGQPPGAKIIPAPVSSVCSQTNALGPTPTAPP- 599  
QY 575 SPPRPQRS 583  
Db 600 AEPALRA 608

## RESULT 15

US-10-145-586-61  
Sequence 61, Application US/10145586  
Publication No. US20030138890A1  
GENERAL INFORMATION:  
APPLICANT: Alexandra Gluckemann, Maria  
APPLICANT: Silos-Santiago, Immaculada  
APPLICANT: M. Galvin, Katherine  
APPLICANT: Weich, Nadine  
APPLICANT: Curtis, Rory A.J.  
APPLICANT: Bandaru, Rajasekhar  
APPLICANT: Kapeller-Liebermann, Rosana  
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,  
TITLE OF INVENTION: HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LEUCINE-RICH  
FILE OF INVENTION: REPEAT FAMILY MEMBERS, AND HUMAN RINGFINGER FAMILY MEMBER  
FILE REFERENCE: 10448-188001  
CURRENT APPLICATION NUMBER: US/10/145,586  
CURRENT FILING DATE: 2002-05-14  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 95  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 61  
LENGTH: 628  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-145-586-61

Query Match 46.7%; Score 1557; DB 14; Length 628;  
Best Local Similarity 54.0%; Pred. No. 46-104;  
Matches 329; Conservative 71; Mismatches 175; Indels 34; Gaps 8;

QY 4 PLLLLLLASGAAA-----CPPLVCQNLSESLSLTCNHRGLFPVNVDRRTVEL 53  
Db 5 PLLLLLPLAPASSPPOSATSPCPRCRCQQTOSLPLSVLCPGAGILFVPPSLDRAAEI 64  
QY 54 RLADNFIOALGPPDRNMGTGLVDLTLSRNATIRIGARAFGOLFSRSLHDGNRLVELGT 113  
Db 65 RLADNFIVASVRRDLANMTGLHLISLSRNTIRHVAAGAFADLRALRLHDGNRLTSLGE 124  
QY 114 GSLRGPNVLOHLISGNOLGRIAPGAFDDFLESLELDLSYNNLRQVPMAGIGAMPAH 173  
Db 125 GOLRLVNLRLHLISNNQALALAGALDCAETLEDLDLSYNNLEQLPMEALGRIGNVT 184  
QY 174 LNLHNLIDALPFGAFOLGOLSRIDLTNSRLATTAPDPLFSR---GRDAEASPAAPVL 229  
Db 185 LGLHNLIVASVPAFSSRLARLDMTSNRLTTPDPLFSRLPLARPRGSPASALVL 244  
QY 230 SFGNPLHNCCELWLRLARPDDLETGASPPGLAGRYFVAVPEGEFCEPPLIARHTQR 289  
Db 245 AFGNPLHNCCELWLRLARDDLEACASPPALGGRYFVAVGEFEFVCEPPEVYTHRSPP 304  
QY 290 LWTLEGGATLRCALGDPATMHWGPDRLVGNSSRARAFPNGLTLEIGVTAGADAGY 349  
Db 305 LAVPAGRPALRCRAVGDPEPRVHVSPQGRLLGNSSRARAFPNGLTLELVTEPDGIF 364  
QY 350 TCITATNPAGETAKAVELRYALPH---CGNSAEGCRPG-----PSDIAASARTAE 399  
Db 365 TCITANAAGEATVAVELVGPVPPQLANSTSCDPPRGGDDALTPSAASAKVADTG 424  
QY 400 EGTLESEPAVQVTEVATSGVSWGCPRPADPVMMFOIQVNSSEDETLIYRIVPASSHIF 459

Db 425 PPT---DRGVQTEHGATAALVQMPDQRPPIGRIMYQIQYNSASSADILVYRMI PAESRSF 481  
QY 460 LKHLVPGADYDLCILALSPAGPSDLTATRLIGCAHFSTLPSPLCHALQAHVLTGTT 519  
Db 482 LITDLSGRTYDLCVLAIVEGSA-TGLTATRPVGARFSTEPALRPGAHAFLLGGTMI 540  
QY 520 VAVGVIVAAALVFTVALVR-----GRGANGRLPLKLSHVOSQTNCGSPPTPKAHPR 574  
Db 541 IALGVIVASVIVFTVFLMRKYKHGGQPPGAKIIPAPVSSVCSQTNALGPTPTAPP- 599  
QY 575 SPPRPQRS 583  
Db 600 AEPALRA 608

Search completed: November 17, 2004, 22:52:27  
Job time : 145 secs

*This Page Blank (uspto)*



```

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-907-794A-245

Query Match      12.8%; Score 425.5; DB 4; Length 713;
Best Local Similarity 25.2%; Pred. No. 1.4e-26;
Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY      1  MAPRLILLASGAA-----CPLPCVCO-----NLSSTLCAHRLFLV 42
DB      4  LVAPLLAMVAGATATVVPVHVPCPCQACQIRPPTSSRTREATTVDCCNDLFLITAV 63
QY      43  PNVVDRTELRLADNFICAGPPFRMTGLVDLTLSRNAITRIGARFGDLESLSLH 102
DB      64  PPAIPATQTLILQSNISIVAVDQELGYLANLTFLDLSQNSFSDARCDFFALPQLLSLH 123
QY      103  LDGNRLVELGTSGRGVNIQHLISNOICRIAPGAFD----- 141
DB      124  LERNQLTRLEDHSPAGLSIQELYNQYRIAPRAFSGLSNLRHLNSNLRALDSR 183
QY      142  -----DFLESLELD 151
DB      184  WFEMLEPULIMTGNKYDALDMNRPPLANLSVLAQMNLRTISDYALEGLQSLSL 243
QY      152  LSYNNLRQVWAGIGAMPALHTTLNLDHNLIDALPPGAF----- 190
DB      244  FYDNLARVRRALAEQVPGJLKFIDLNKNPIQVGGPFAMNMLHKLGLNMMELVSDK 303
QY      191  ----OLGQSLRLDITSN-RLATLAPDPLFS-----RGRDAKSPAPLV 228
DB      304  FALVNLPELTETLDITNNPRLSFIPRAFHMLPQMETIMNNALSALHQCVESLPIQOE 363
QY      229  LSFSGNHLNHCCELLMWR-----RLARPDLEFCASPPGAGRYFAVPEGEFS--CEP 280
DB      364  VGLHGNPRCCCVIRMANAIVTRVRFTEPOS-TICAEPPDIQRIPVAREVPREMTDCHP 422
QY      281  PLIAR-HTORLMLVEGORATLRCRALGDPAPFTMHWGDD-RLVGNSS--RARAPENGTL 336
DB      423  LISRSPPSLSQVASGESMVLHCRALAPBEIYWTVPAGRLTPAHAGRRYRVPEGTL 482
QY      337  EIGVTGAGDAGGYCIATNPGEATAYELAVLPHQGNSSAGCGRPESDIAASRTA 396
DB      483  ELRVTABEAGLYTCVQNLVAGADTKTVSVV-----GRALLQPGRD----- 524
QY      397  AEGGTLESEBPAVOVTEVATSGLVSW--GGRPADPVMMFOIQVNSEDETLTYRIYPA 454
DB      525  -EGGS-LR---LVQETHPHILISVTPNTYSTNLTWSSASLSKGGATALARL-ER 577
QY      455  SSHHFLKHLVPGADYDCLALSPAGPSDLTATRLIGCAHFSTLPASPLCHALQAHVL 514
DB      578  GTHSYNITRLIQATEVYMACIQV-----APADAHQLQACVWARTKEATS-CHR----- 623
```

```

QY      515  GGLTVAVG---GVVVAALLVFTVALLVGRGA--GNGRLPLKTSIVSQITNGGSPPTPK 569
DB      624  -----ALGDRPGLI--ATLALAVILLAGLAHLGTGQ-PRK-----GVGGRRPLPP 667
QY      570  A 570
DB      668  A 668

RESULT 2
US-09-905-125A-245
; Sequence 245, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltzen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-125A-245

Query Match      12.8%; Score 425.5; DB 4; Length 713;
Best Local Similarity 25.2%; Pred. No. 1.4e-26;
Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPRLLLLLASGAAA-----CPRLVCQ-----NLSSSLTLCARGLLHV 42
DB 4 LVAPLLAVAGATATVPVPMVPCPCQACQIRPMYTPRSSYREATTVDNDLPLTAV 63
QY 43 PPNVDRRTVELRLADNFIQALGPPDFRNMGTGLVDLTLSRNATRIGARAFGDELSRSLH 102
DB 64 PRALPRAGTQTLILQNSIVRVDSSELGYLANLTLELDSQNSFSDARDCCFHALLPQLSLH 123
QY 103 LDGNRIVELGTGSLRGPVNLQHLILSNGOLGRIAPGAFD----- 141
DB 124 LERNQUTRLDEHDFAGLASIQELYNHNLQYRIAPRAFSGLSNLRHLNSNLRALDSR 183
QY 142 -----DPLESEDD 151
DB 184 WERMENLEILMIGNKVDAILDMNFRPLANISLVLAGMNLREISDYALEGLQSLSEIS 243
QY 152 LSYNNLRQVPMWAGIAMPALHTLNDNLIDALPQAPAF----- 190
DB 244 FYDNOGLARVPRRLLEGVPGKPLDKNPLQRVGPQDFANMLHKLGLANNEELVSTDK 303
QY 191 ----QIGQLSRDLTISN-RIATLAPDPLFS-----RGRDAEASAPPLV 228
DB 304 FALVNLPELTKLDTNNPRLSFHPRAFHNLPMETMLNNNALSAHQVLESLENLGE 363
QY 229 LSFSGPRLNCNCLLWLR-----RLARPDULETCSFPGLAGRYWAVPEGEFS--CEP 280
DB 364 VGLHNPRIKDCVIRANAMATGRVRIEPOS--TLCAEPDQLRLPVREVPFEMTDHCHP 422
QY 281 PLIAR-HTQRLWVLESGORATLRCRAGDPAFTWGWGPPD-RLVGNSS--RARAPNGL 336
DB 423 LISPRFPFSLQVASESMVLRKALAEDEPELWVTPAGLRLTPAHAGRRTRVPEGL 482
QY 337 ELGVAGAGAGYTCIATNPAGEATARVELRYLALPHGNSASABGGRPGPSDIASARTA 396
DB 483 ELRRVTAEBRAGLYTCVAGNMGADTKTVGVVV-----GRALLQPERD----- 524
QY 397 AGEGETLESEPAVOVTEVATSGLVSW--GPGRPADPVMMFOIQYNSSDEDTLYIYIVA 454
DB 525 -BGQG-LR---LRVQETHPHYHLLSWTPRPNTVSTNLTWSSASSLRGGGATALARTL-BR 577
QY 455 SSHHFLKLHLPVGDVLDLCLALSPAGSDLTATRLLCAPHSSTLPASPLCAHQANVL 514
DB 578 GHHSVITRLQLATEWACLQV-----AFADAHQLQALCWAARTKENS-CHR----- 623
QY 515 GGTLLVAVG---GVVVAALLVFTVALLVGRGA--GNGLPLKLSHVQSQTNGGSPPTPK 569
DB 624 -----ALGDRPGLI--ATLALAVILLAAGLAAHGTGQ-PRK-----GVGGRPLRP 667
QY 570 A 570
DB 668 A 668
```

```

RESULT 3
US-09-902-775A-245
; Sequence 245, Application US/09902775A
```

```

; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Bostein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurley, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-775A-245
```



Query Match 12.8%; Score 425.5; DB 4; Length 713;  
 Best Local Similarity 25.2%; Pred. No. 1.4e-26;  
 Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPPLILLILASGAAA-----CPLPCVCO-----NLSLSLTLCAHRLFLV 42  
 DB 4 LVAPLLAWAGATATVPVPMWHPVPCPCQACQIRPMYTPRSSYRBAITVDGNDLFLTAV 63  
 QY 43 PPNVDRRTVELRLADNFIQALGPPDFRNMGTGLVDLTSRAITRIGARAGDELSRLSH 102  
 DB 64 PPALPACTQTLILLOSNSIVAVDQSELGYLANLTLELDSQNSFSFARDCCDFHALPOLLSH 123  
 QY 103 LDGNRLVELGTSLRGVNLQHLILSGNQLGRIAPGAFD----- 141  
 DB 124 LEENQLTRLEHDSFAGLASLQELLYLNHNLQYRIAPRAFSGLSNLRHLNSNLRADSR 163  
 QY 142 -----DFLESLEDD 151  
 DB 184 WPEMLPVLITLMTGKNKVDAILDKNFRPLANLSVLAKNNLREISDYALEGLQSLSL 243  
 QY 152 LSTVNNLRQVWAGIGAMPALHTLNDNLIDALPRGAF----- 190  
 DB 244 FYDNLARVRPRALEQVPGKFLDNLNPLQRVGPGDFAMMLKELGLNMEELVSIDK 303  
 QY 191 -----OLGOLSRDLTSLN-RLATLAPDPLFS-----RGRDAEASPAFLV 228  
 DB 304 FALVNLPELTKLDTNNPRLSFIHPRAFPHLPQMETLMMNNLSALHQGVESLPLQE 363  
 QY 229 LSPSGNPLHCNCELLMWR-----RLARPDDLETCASPGIAGRYFAVPEGES--CEP 280  
 DB 364 VGLHGNIRCDQVIRMANATGTRVRFTLEPOS-TLCAEPPLQRLPVRVPEFRETWDHCLP 422  
 QY 281 PLIAR-HTQGLWLEEGORATLRGALGDPAPMTMMWCPDD-RLVGNSS--RAAFEPGTL 356  
 DB 423 LISPRSPPLQVASSGSMVLCRCLAEPEPELYWTPAGRLTPAHAGRRYRYPGTL 482  
 QY 337 BIGVTGAGDAGGYTCIATNPAGEATARELEVLALPHGNSASAGRPGPSDIASARTA 396  
 DB 483 ELKRVTLAEKGLTYCQONLVGADTKTVSVV-----GRALLQPGD----- 524  
 QY 397 AEGGTLESEPAVQVETVATSGLVSW-GEGRPADPVMWFQIYNSEDETLIYRIVPA 454  
 DB 525 -EGOG-LE---LAVQETHPYHILSWTPENTVSTMLTSSASLSRGQATALARL-PR 577  
 QY 455 SSHHPLKLHVPAGDYDLCLLSPAGPSDLTTRLLGCAHFTLPAAPLCHALQHV 514  
 DB 578 GTHSYNITRLLQATEYACQOV-----APADHQLQACVMARKETKS--CHR----- 623  
 QY 515 GGLTVAVG--GVIVALLVFTVALIVRGGA--GNGRLPLKLSHYOSQTNQSPSPTPK 569  
 DB 624 -----ALGDRPGLI--ALLANLVLLAAGLAHLGTGQ-PRK-----GVGGRRLPLP 667  
 QY 570 A 570  
 DB 668 A 668

RESULT 4  
 US-09-906-700-245

; Sequence 245; Application US/09906700  
 ; Patent No. 6723535  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genentech, Inc.  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, A.  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, Christopher J.  
 APPLICANT: Guiney, Austin L.  
 APPLICANT: Hillan, Kenneth, J.  
 APPLICANT: Kljavin, Ivar J.  
 APPLICANT: Mather, Jennie P.  
 APPLICANT: Pan, James  
 APPLICANT: Peoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tuma, Daniel  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William, I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 TITLE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: 10466-14  
 CURRENT APPLICATION NUMBER: US/09/906,700  
 CURRENT FILING DATE: 2000-09-18  
 PRIOR APPLICATION NUMBER: PCT/US00/04414  
 PRIOR FILING DATE: 2000-02-22  
 PRIOR APPLICATION NUMBER: US 60/143,048  
 PRIOR FILING DATE: 1999-07-07  
 PRIOR APPLICATION NUMBER: US 60/145,698  
 PRIOR FILING DATE: 1999-07-26  
 PRIOR APPLICATION NUMBER: US 60/146,222  
 PRIOR FILING DATE: 1999-07-28  
 PRIOR APPLICATION NUMBER: PCT/US99/20594  
 PRIOR FILING DATE: 1999-09-08  
 PRIOR APPLICATION NUMBER: PCT/US99/20944  
 PRIOR FILING DATE: 1999-09-13  
 PRIOR APPLICATION NUMBER: PCT/US99/21090  
 PRIOR FILING DATE: 1999-09-15  
 PRIOR APPLICATION NUMBER: PCT/US99/21547  
 PRIOR FILING DATE: 1999-09-15  
 PRIOR APPLICATION NUMBER: PCT/US99/23089  
 PRIOR FILING DATE: 1999-10-05  
 PRIOR APPLICATION NUMBER: PCT/US99/28214  
 PRIOR FILING DATE: 1999-11-29  
 PRIOR APPLICATION NUMBER: PCT/US99/28313  
 PRIOR FILING DATE: 1999-11-30  
 PRIOR APPLICATION NUMBER: PCT/US99/28564  
 PRIOR FILING DATE: 1999-12-02  
 PRIOR APPLICATION NUMBER: PCT/US99/28565  
 PRIOR FILING DATE: 1999-12-02  
 PRIOR APPLICATION NUMBER: PCT/US99/30095  
 PRIOR FILING DATE: 1999-12-16  
 PRIOR APPLICATION NUMBER: PCT/US99/30911  
 PRIOR FILING DATE: 1999-12-20  
 PRIOR APPLICATION NUMBER: PCT/US99/30999  
 PRIOR FILING DATE: 1999-12-20  
 PRIOR APPLICATION NUMBER: PCT/US00/00219  
 PRIOR FILING DATE: 2000-01-05  
 NUMBER OF SEQ ID NOS: 423  
 SEQ ID NO 245  
 LENGTH: 713  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-09-906-700-245

Query Match 12.8%; Score 425.5; DB 4; Length 713;  
 Best Local Similarity 25.2%; Pred. No. 1.4e-26;  
 Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPPLILLILASGAAA-----CPLPCVCO-----NLSLSLTLCAHRLFLV 42  
 DB 4 LVAPLLAWAGATATVPVPMWHPVPCPCQACQIRPMYTPRSSYRBAITVDGNDLFLTAV 63  
 QY 43 PPNVDRRTVELRLADNFIQALGPPDFRNMGTGLVDLTSRAITRIGARAGDELSRLSH 102  
 DB 64 PPALPACTQTLILLOSNSIVAVDQSELGYLANLTLELDSQNSFSFARDCCDFHALPOLLSH 123  
 QY 103 LDGNRLVELGTSLRGVNLQHLILSGNQLGRIAPGAFD----- 141

```

Db      124 LEENQTLRLBDHSPAGLASLOEYLYLNHNOYRIAPRAFSGLSNLLRLHNSNLLRAIDSR 183
Qy      142 -----DPLESLEDJD 151
Db      184 WFEMLPNLEILMIGNKVDAILDMNFRPLANLSVLAGMNLREISDYALEGLQSLESLS 243.
Qy      152 LSYNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF-----190
Db      244 FYDNLARVPRALQVPGELKFLDLNKNPLQRVGPGFAMMLHUKELGLNMEELVSDK 303
Qy      191 -----QLGQSLRLDLSN-RLATLADPLFS-----RGDAASPAPLY 228
Db      304 FALVNPETLTKLITNNPRLSFTHPRFHHLPQWETMLNNNALSAHQCVESLPLQGE 363
Qy      229 LSFSGNPLHCNCELLMWR-----RLARPDDLETCASPPGLAGRYFMAVEGEFS--CEP 280
Db      364 VGHGNPFRCDCTIRMANNAITGTVRFLEPOS-TLCAEPPDQRLPVREVPFRMTDCLP 422
Qy      281 PLIAR-HTQRLWVLEGGRATLRCRAGDPATMHWGPPD-RLVGNSS--RADAFFNGTL 336
Db      423 LISPSPFPPLQVASSGSWVLHCRALAEPEPEIYWTBAGRLTPHAGRRVYVEGTL 482
Qy      337 EIGVTAAGDAGGTCTATNAGATATARVELRYVALPHGNSSAEGGRPPGSDIAPART 396
Db      483 ELRRVTAEEAGLTVCAQNLVGADITVSVV-----GRALDQPGSD-----524
Qy      397 AEGEGTLESEPAVQVETVTSGLVSW--GPRPADPVMWFQIQVNSSEDETLIYRIVPA 454
Db      525 -EGQ-LE-----LRVQETHYHLLLSWVPPTVSTNLTSSASSLSRGQATILARL-PR 577
Qy      455 SSHEFLKLHVPADYDCLLALSPAGSDLTATRLGCAHFTLPASPPLCHALQAVL 514
Db      578 GTHSYNITRLQTEYEWACLOV-----AFADAHQQLACVMARTEANS-CHR-----623
Qy      515 GGLTLTAVG---GVVVAALVVFVALLVGRGA--GNGRLPLKLSHVQSCITNGSPPTPK 569
Db      624 -----ALGDRPLI--AILALAVLLIAGLAHLGTQ-PRK-----GVGGRPLDP 667
Qy      570 A 570
Db      668 A 668

```

```

RESULT 5
US-09-903-603A-245
; Sequence 245, Application US/09903603A
; Patent No. 6767995
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gueney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mathey, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey

```

```

; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: GNE.1618P2C12
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; TYPE: PRT
; LENGTH: 713
; ORGANISM: Homo Sapien
; US-09-903-603A-245

```

```

Query Match      12.8%; Score 425.5; DB 4; Length 713;
Best Local Similarity 25.2%; Pred. No. 1,4e-26;
Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

```

```

Qy      1 MAPILLILLASGAAA-----CPDPVCQ-----NLSSTLCAHRLGLTV 42
Db      4 LVAPLLAWAGATATVPVPMHVPCPPQCAQIRPWYTPRSSYREATVDCNDLFTLV 63
Qy      43 PENVRRTVELRLANFTIALGPPDFRMNTGLVDLTLSRNAITRIGARFGDLESLSRH 102
Db      64 PPAIPAGTQTLLONSIVRVQSELGYIANLTEDLSGNSGSDARDCCFHALLPOLLISLH 123
Qy      103 LDGNRLVELGTSLGPNVLQHLISGQGLRIAGAPD-----141
Db      124 LEENQTLRLBDHSPAGLASLOEYLYLNHNOYRIAPRAFSGLSNLLRLHNSNLLRAIDSR 183
Qy      142 -----DPLESLEDJD 151
Db      184 WFEMLPNLEILMIGNKVDAILDMNFRPLANLSVLAGMNLREISDYALEGLQSLESLS 243
Qy      152 LSYNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF-----190
Db      244 FYDNLARVPRALQVPGELKFLDLNKNPLQRVGPGFAMMLHUKELGLNMEELVSDK 303
Qy      191 -----QLGQSLRLDLSN-RLATLADPLFS-----RGDAASPAPLY 228

```

```

Db      304 FALVNLPELTKLDITNNPRLSFHPRAFHLLPQMETILNNNALSAHQOVSLEPNLOE 363
QY      229 LSSGNPLHNCNCLMLR-----RLARPDLLETQSPGLAGRYFWAVEGEFSC----- 278
Db      364 VGHGNPIRDCVIRMANATGTRVRIEPOS-TLCAEPDLOQLPVREVPFRMTDCHLP 422
QY      281 PLIAR-HTORLWLEQORATLRCRAGDPPATMHWGPPDD-RLVGNSS--RARAFTNGTL 336
Db      423 LIIPRFPPELOVASGESWVLHCRALAEPEIYWTPLAGRLPLPAHAGRRVYVEGTL 482
QY      337 EIGVTAGAGAGYTCIATNPAGEATYARVELVIALPHGNSSAEGRPESDIAASART 396
Db      483 ELRRVTAEEAGLYTCAQNLVGADTKTVSVV-----GRALLQPGPD----- 524
QY      397 ABEGTLESPPAQTVEVTATSGVW--GPRPADPVMWFOIQNNSSEDETLIYIVPA 454
Db      525 -EEQG-LE-----LRVETHPYHLLSWPPTVSTNLVSSASSLRGGATLARI-PR 577
QY      455 SSHHFLKHLVPADVLCIALSPAAGPSDLTATRLGCAHFTLPASPLCHALQAVL 514
Db      578 GTHSYNTRLLOATEYKACIQV-----AFADAHQOLACVMARKTEATS-CHR----- 623
QY      515 GGTITVAVG--GVLVAALLVFTVALIVRGGA--GNGRLPLKSHVQSQTNGPSPETPK 569
Db      624 -----ALGDRPGLI--AIIALAVLLAAGLAHLGTGQ--PRK-----GVGRRPLLP 667
QY      570 A 570
Db      668 A 668

```

## RESULT 6

```

US-09-063-950-2
; Sequence 2, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-063-950-2

```

```

Query Match      12.4%; Score 415; DB 3; Length 673;
Best Local Similarity 26.2%; Pred. No. 9.5e-26;
Matches 176; Conservative 69; Mismatches 280; Indels 148; Gaps 22;

QY      1 MAPPLILLILASGAACPILPCVCONLSSTLCAHRLGLFVPPNVDRTVELELADNFI 60
Db      8 LIRPLILLALGPVGQSPSGQC-----SQPQVFTAGSTTVPRDVPDVGIVFENGFI 64
QY      61 QALGPPFRNMGTGLVLTLSRNATIRIGARAFGDLSEIRSLHLDGNRLVETGSLGPGV 120
Db      65 TMLDAGSFAGLPGLQILDLSQNOIASLPSGVFOPLANISNIDLTPANRIHEITNETFFGLR 124
QY      121 NIQHLILSGNQLGRIAPGAPDFLESLEDLSYNNLRQVWAGIGAMPALHTNLNL 180
Db      125 RLSELYIGKNIRIHTQGAFT-LDRLLLEKLQNEIRALPPL---RLPRLILLDLSHNS 180
QY      181 IDALPP-----GAFAQLGQLSRLLDLSN----- 203
Db      181 LIALEPGILDTANVEALRIAGLGLQQLDEGLFSRLRNHLDDVDNQLEVPVYIRGLRG 240
QY      204 -----RLATLADPLFS-----RGDAEASAPPL-----VLSFGSNPL 236
Db      241 LTRRLAGNTRIAQIRBEDLAGLAALQELDVSNLSIALPGLDLSGLPRLRLTAAARNPF 300

```

```

QY      237 HCNCELLMR-----RLARPDLLETQSPGLAGRYFWAVEGEFSC----- 278
Db      301 NCVCPLSMGPNVRESHYTLASPEETR-CHFPKNAGRLLLELDVADFGCPATTTATVP 359
QY      279 -----EPPLIARHTORLWLEQORATLRCRAGDPPATMHWGPPDDRLVGNSSARAF 331
Db      360 TTRPVREPTLASSIAPFTMLSPATAT--EAPSPSPSTAPPTVGVPPQ-----PDQCP 410
QY      332 PNCITLIGVTGADGAGYTCIATNPAGEATYARVELVIALPHGNSSAEGRPESDIAA 391
Db      411 PSTCLNGGTCIHGTRHHLALC--PEFTGLVCE-----SOMGQTRSPSPVVP 458
QY      392 SARTAEGEGLSEPAVQTEVTATSGVWPGRPADPVMWFOIQN--SSEDETLIY 449
Db      459 RPRSL-----TLGIEP-----VSPSTSLAVGLQRYLGSSVQ-LRSLRLTYRLSGPDRKLV 510
QY      450 RIVPASHHFLKHLVPADVLCIALSPAAGPSDLTATRLGCAHFTLPASPLCHAL 509
Db      511 LRLPASLAETVYQLRPNATVSVCPMLPGRVPEGEBA-----CGEANTPPVHSHAP 565
QY      510 QAHVIGGTLTV-----AVGVVAALLVFTVALIV-RRG-----AGNGRLPLKLS 554
Db      566 VTQAREGNPLLIAPLAAVLLAALAAVGAAYCVRRGRAMAAAADQXGVGPAGPLELE 625
QY      555 HVQSQTNGPSPPT 567
Db      626 GVKVPLEPGRKAT 638

```

## RESULT 7

```

US-09-907-794A-292
; Sequence 292, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Etkin, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvarolf, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28

```

```
/ PRIOR APPLICATION NUMBER: PCT/US99/20594
/ PRIOR FILING DATE: 1999-09-08
/ PRIOR APPLICATION NUMBER: PCT/US99/20944
/ PRIOR FILING DATE: 1999-09-13
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/21547
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/23089
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: PCT/US99/28214
/ PRIOR FILING DATE: 1999-11-29
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: 1999-11-30
/ PRIOR APPLICATION NUMBER: PCT/US99/28564
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/28565
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: PCT/US99/30911
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US99/30999
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US00/00219
/ NUMBER OF SEQ ID NOS: 423
/ SEQ ID NO 292
/ LENGTH: 640
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-09-907-794A-292
```

```
Query Match      12.2%; Score 407.5; DB 4; Length 640;
Best Local Similarity 26.8%; Pred. No. 3.7e-25;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;
```

```
QY 4 PHLILIA-----SG--AAACPLPCVCNTSESISTICANGLLFVPPNDRRTVEIR 54
DB 25 PLIVVLLAIQLVAVAGLVRAQTCPSCVCSN--QFSKVICVRKMLEVEPDGISTNRLIN 82
QY 55 LADNFQALGPFRFRMTGVDLTLSRNATIRGAAGFDLBSRLHMDGNFLVELGTG 114
DB 83 LHENQIITKXNSFKLRHLEITLQSRNHRITIEIGFNGLANLTLEFDNLTTPNG 142
QY 115 SLRGPVNLQHLISGN-----QUGRIAPGAPDDFESLSD 149
DB 143 AFVYLSKLEIKWRNNPISISYARNRIPSLRLDLGELKRLSYISGAFBS-LSNLRKY 201
QY 150 LDI SYNNLRQVPAWAGIGAMPALHTLNLIDALPPGAFAOLGQLSRU----- 198
DB 202 LNLAMCNLRKRP--NLTPILKIDELDLSCNHLISAIRGSGFQGLMHLQKLMVIGSQIYTE 259
QY 199 -----DLTSNLTATAPPLFSRGRDASAPVYLSGSGNPLHNCCELLMW 245
DB 260 RMAFDNLQSLVEINLAHNL-TLLPDLFTPLHHLER-----IHLHNPANCNCIILWL 312
QY 246 R---RLAPDDLETCA---SPGLAGRYFAVYBEGEFSCEPLIAHNTQRLMWLESGRAT 299
DB 313 SWMKIDMAISNTACRCARCTPPNLKRYIGELDQNYFTYARIVYPPADLVNTBSMALE 372
QY 300 LKCRALGDPAPTMWVGPPDRILVGNSS---RARAFNGTLEIGVTGAGAGGYCTIATNP 356
DB 373 LKCRALGDPAPTMWVGPPDRILVGNSS---RARAFNGTLEIGVTGAGAGGYCTIATNP 356
QY 357 AGEATRVELRYAL-----PHGNSAABEGRPSPSI---AASANT 395
DB 432 VGNITASATLNTYATATTPPSFYSTVYETVEMPSQDEARTDNNVGPVVDMETNVT 491
QY 396 AAEGBGLESPRAVQYETVATSGLVSWPG 426
DB 492 SLTPQSTRSTKFTTIPVDINSGL-----PG 518
```

```
RESULT 8
US-09-905-125A-292
/ Sequence 292, Application US/09905125A
/ Patent No. 6664376
/ GENERAL INFORMATION:
/ APPLICANT: Genentech, Inc.
/ APPLICANT: Ashkenazi, Avi
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferreira, Napoleone
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerlitsen, Mary E.
/ APPLICANT: Goddard, A.
/ APPLICANT: Grimaldi, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillen, Kenneth, J.
/ APPLICANT: Kijavlin, Ivar J.
/ APPLICANT: Mather, Jennie P.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Thomas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William, I.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: 10466-14
/ CURRENT APPLICATION NUMBER: US/09/905,125A
/ PRIOR FILING DATE: 2001-07-12
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: 2000-02-22
/ PRIOR APPLICATION NUMBER: US 60/143,048
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: US 60/145,698
/ PRIOR FILING DATE: 1999-07-26
/ PRIOR APPLICATION NUMBER: US 60/146,222
/ PRIOR FILING DATE: 1999-07-28
/ PRIOR APPLICATION NUMBER: PCT/US99/20594
/ PRIOR FILING DATE: 1999-09-08
/ PRIOR APPLICATION NUMBER: PCT/US99/20944
/ PRIOR FILING DATE: 1999-09-13
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/21547
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/23089
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: PCT/US99/28214
/ PRIOR FILING DATE: 1999-11-29
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: 1999-11-30
/ PRIOR APPLICATION NUMBER: PCT/US99/28564
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/28565
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: PCT/US99/30911
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US99/30999
/ PRIOR APPLICATION NUMBER: PCT/US00/00219
/ NUMBER OF SEQ ID NOS: 423
/ SEQ ID NO 292
/ LENGTH: 640
```

```

; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-125A-292

Query Match
Best Local Similarity 26.8%; Score 407.5; DB 4; Length 640;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

QY 4 PLILILIA-----SG---AAACPLPCVCONLSSESLTCAHKGGLFVPPNVDRRYEELR 54
   |||:|||||
Db 25 PLIVVILAIQLVLVAGVRAQTCPVSCSN--QFSKVICVRKNLREVDPGISNTNRLIN 82
   |||:|||||
QY 55 LADNFIQALGPPPRNMTGVDLTLSNATRIIGARAFGDLRSIHLIDGNLVEGTG 114
   |||:|||||
Db 83 LHENQIITIKVNSFKHLRHEILQLSRNHRITIEIGAFNGLANLNTLELFDNLTTIPNG 142
   |||:|||||
QY 115 SLRGPNVLOHLISGN-----OLGRIAPGAFDDFLESLED 149
   |||:|||||
Db 143 AFVYLSKLKELWLRNNPTESI PSYAFNRIPSLRRLDGLKRLSYISEGAFEG-TSNLRY 201
   |||:|||||
QY 150 LDISYNNLRQVPMAGICAMPALHTLNDHLIDALPGAFQQLGQLSRL----- 198
   |||:|||||
Db 202 LNLAMCNLRIRP--NLPLRLKLDLDSGNHLSAIRGSPQGLMHLQKLMWISQIQVIE 259
   |||:|||||
QY 199 -----DLTSNRLATLAPDLFSRGRDAEASPAVLVSFGNPLHNCCELLML 245
   |||:|||||
Db 260 RNAPFNLQSLVEINLAHNNL-TLLPHDLFTPLHLHER-----IHLHNPWNCNDILML 312
   |||:|||||
QY 246 R---RLARPDDLETCA---SPPGIAGRYFWAVBGEFSCEPPLIARHTORLWLEGGORAT 299
   |||:|||||
Db 313 SWMIKQMAPENTACCAACNTPPNLRKRYIGELDQNYTTCAPVIVBPPADLNTBGMAAE 372
   |||:|||||
QY 300 LRCRALGDPAPTMHVGPDRLVGNSS--RARAFNPGTIRIGVTGAGDAGTTCIATNP 356
   |||:|||||
Db 373 LKCRRA-STSLTYSWITPNGVMTGAYKVIAYLSGTLNFTNVTQDDGMTCWMSNS 431
   |||:|||||
QY 357 AGETARVELRYVAL-----PHGNSAEGGRPGPSDI---AASART 395
   |||:|||||
Db 432 VGNTTATATNTVNTATTTPPSYSTVETVETMEPSQDARITDNNVGTPTPVVDETTNVT 491
   |||:|||||
QY 396 AABEGTLESEPAVQVTEVTATSGLVSMGPG 426
   |||:|||||
Db 492 SLTPQSTRSTEKPTTIVTDINSGL---PG 518
   |||:|||||

RESULT 9
US-09-902-775A-292
; Sequence 292, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Amkenazal, Avi
; APPLICANT: Bostein, David
; APPLICANT: Deenoysers, Luc
; APPLICANT: Eaton, Nan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
```

```

; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902, 775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 292
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-775A-292

Query Match
Best Local Similarity 26.8%; Score 407.5; DB 4; Length 640;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

QY 4 PLILILIA-----SG---AAACPLPCVCONLSSESLTCAHKGGLFVPPNVDRRYEELR 54
   |||:|||||
Db 25 PLIVVILAIQLVLVAGVRAQTCPVSCSN--QFSKVICVRKNLREVDPGISNTNRLIN 82
   |||:|||||
QY 55 LADNFIQALGPPPRNMTGVDLTLSNATRIIGARAFGDLRSIHLIDGNLVEGTG 114
   |||:|||||
Db 83 LHENQIITIKVNSFKHLRHEILQLSRNHRITIEIGAFNGLANLNTLELFDNLTTIPNG 142
   |||:|||||
QY 115 SLRGPNVLOHLISGN-----OLGRIAPGAFDDFLESLED 149
   |||:|||||
Db 143 AFVYLSKLKELWLRNNPTESI PSYAFNRIPSLRRLDGLKRLSYISEGAFEG-TSNLRY 201
   |||:|||||
QY 150 LDISYNNLRQVPMAGICAMPALHTLNDHLIDALPGAFQQLGQLSRL----- 198
   |||:|||||
Db 202 LNLAMCNLRIRP--NLPLRLKLDLDSGNHLSAIRGSPQGLMHLQKLMWISQIQVIE 259
   |||:|||||
QY 199 -----DLTSNRLATLAPDLFSRGRDAEASPAVLVSFGNPLHNCCELLML 245
   |||:|||||
Db 260 RNAPFNLQSLVEINLAHNNL-TLLPHDLFTPLHLHER-----IHLHNPWNCNDILML 312
   |||:|||||
QY 246 R---RLARPDDLETCA---SPPGIAGRYFWAVBGEFSCEPPLIARHTORLWLEGGORAT 299
   |||:|||||
```

Db 313 SWIKMAPSNTACCAKRCNTPNVLKGRYIGELDONFTTCAPVIVPPADLNTTEGMAAE 372  
Qy 300 LRCRALGDPAPTMHWWGPDRLVGNSS---RARAFNGTLEIGVTGAGAGGTCIATNP 356  
Db 373 LKCRRA-STSLTSVSMITPMTGVTMTHGAYKVRIVLSDGTLNFTNVTVOOTGMVTCVNS 431  
Qy 357 AGEATARVELRYLAL-----PHGNSSAEGRPGPSDI---AASART 395  
Db 432 VGNTTASATLNTVTAATTPFSYSTVETVETMERESQDEARTDNNVGPFPVDMETNVT 491  
Qy 396 AAEGETLESEPAVQVTEVTATSGLVSWGPG 426  
Db 492 SLTPGSTSTKTEFTIIPVTDINSGI---PG 518

RESULT 10  
US-09-906-700-292  
Sequence 292, Application US/09906700  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/906,700  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 292  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-906-700-292

Query Match 12.2%; Score 407.5; DB 4; Length 640;  
Best Local Similarity 26.8%; Pred. No. 3.7e-25;  
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

Qy 4 PLILLILA-----SG---AAACPLPCVCONLSSTICARGLLFPVNDRTVEIR 54  
Db 25 PLIVVLLALQLLVAGLVRAQTCPSVCSN--QFSKVI CVRKNLREVPDGI STNRLLN 82  
Qy 55 LADNFIQALGPPDFRNMTGLVLTLSRNAITIGARAFGDLISRLHLDGRVLVLGIG 114  
Db 83 LHENQIILIKVNSFPLRHLEILQLSRNIRITIEGAFGLNANLLELFDRLTLTIPNG 142  
Qy 115 SLRGPVNIQHLILSGN-----QGRIRAPGAFDDFLESILD 149  
Db 143 ARVYLSKLKEMLKRNPIESISYAFNRIRPSRLDLGLKLTYSIGAFBS-LSNLX 201  
Qy 150 LDIYNNIRQVPMAGIGAMPALHTLNLDNLIDALPPGAFOIGQLSRU----- 198  
Db 202 LNLAMCNIREIP--NLTPLIKIDEIDLSGNHLSAIRPGSFQGLMHLQKLMIGIQOIVIE 259  
Qy 199 -----DTSNRLATLADPLFRSGRDAASAPVLSGPNLHNCCELLW 245  
Db 260 RNAFNLQSLVEINLAHNNL-TLPPDLFTPLHHER-----IHNNPNCNCIDILW 312  
Qy 246 R---RLARPDDLETC---SPPLAGRYPMWAVEGFSCEPPLIARHTQRWLWLEQGRAT 299  
Db 313 SWIKMAPSNTACCAKRCNTPNVLKGRYIGELDONFTTCAPVIVPPADLNTTEGMAAE 372  
Qy 300 LRCRALGDPAPTMHWWGPDRLVGNSS---RARAFNGTLEIGVTGAGAGGTCIATNP 356  
Db 373 LKCRRA-STSLTSVSMITPMTGVTMTHGAYKVRIVLSDGTLNFTNVTVOOTGMVTCVNS 431  
Qy 357 AGEATARVELRYLAL-----PHGNSSAEGRPGPSDI---AASART 395  
Db 432 VGNTTASATLNTVTAATTPFSYSTVETVETMERESQDEARTDNNVGPFPVDMETNVT 491  
Qy 396 AAEGETLESEPAVQVTEVTATSGLVSWGPG 426  
Db 492 SLTPGSTSTKTEFTIIPVTDINSGI---PG 518

RESULT 11  
US-10-140-002-368  
Sequence 368, Application US/10140002  
Patent No. 6725730  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey

```

? APPLICANT: Godowski, Paul J.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Sherwood, Steven
? APPLICANT: Smith, Victoria
? APPLICANT: Stewart, Timothy A.
? APPLICANT: Tunnas, Daniel
? APPLICANT: Watanabe, Colin K
? APPLICANT: Wood, William
? APPLICANT: Zhang, Zemin
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? TITLE OF INVENTION: ACIDS ENCODING THE SAME
? FILE REFERENCE: P3330R1C59
? CURRENT APPLICATION NUMBER: US/10/140,002
? CURRENT FILING DATE: 2002-05-06
? Prior Application removed - See Palm or File Wrapper
? NUMBER OF SEQ ID NOS: 550
? SEQ ID NO 368
? LENGTH: 640
? TYPE: PRT
? ORGANISM: Homo Sapien
?
US-10-140-002-368

```

```

Query Match      12.2%; Score 407.5; DB 4; Length 640;
Best Local Similarity 26.8%; Pred. No. 3.7e-25;
Matches 137; Conservative 66; Mismatches 20; Indels 105; Gaps 16

QY          4 PLALLLLA-----SG---AAACPLPCVCQNLSBSLSTLCARHGLLFYPVNVDRRTVELR 54
Dd          25 PLLVVLLALQLLVVAAGLVRAQTGTCSCSN--QRSKVIQVRKNLRBVPDGISTNTRLLN 82
QY          55 LADNFIALGPDPFRNMTCGLVDLTISRALTIRIGARAQGDLESLSHLIDGRNLVELGTG 114
Dd          83 LHENOQIOTIKNSRKHLRHLEILQLSRNHITIBIGAENGMLANTLLELPDRRLTTIPNG 142
QY          115 SLRGPVNLOHTILSGN-----OLGRIAGAQQDFLESIED 149
Dd          143 AFVLISKLEKLWLNRPNIPIESIPSAFNRIPSLRRLDLGELKRSLTISEGAPEG-LSNTRY 201
QY          150 LDLSYNNLRQVPMAGIGAMPALHTLMDLHNLDALPGAFAQLGOLSRL----- 198
Dd          202 LNLAMCNLRERP--NLTPRLKLDLEDLDSGNHLSAIRPSPFQGLMHLQKLMIMIQOIIQYLE 259
QY          199 -----DITSNRLATLAPDPLFSRCRDAAEPALVLSFGSNGPLHCNCILLWL 245
Dd          260 RNAFDNIQSIVLEINLAHNNTLTLPHDFTTPLHILER-----IHJHNPNMNCDDIIML 312
QY          246 R---RLRRPDDELTA---SPPGLAGRYEMAVPEGEFCEPRLIAHRHQRLVNLVEGQRAT 299
Dd          313 SWTKIDMAESTTACCARCNPTRPNLKRGTYIGEILDQVFCTYAVALIEPPADLVETGMABE 372
QY          300 LRCRALGDPARTMEWVGDDRLVGNSS---PARAFPNGTLEIGTAGDAGGYTCIANTP 356
Dd          373 LKCRNA-STSLTSVGSITENGTVMTHGAKVRIAVALSDGLTNFTNTVTDGTGYTCMVSNS 431
QY          357 AGEATARELKVLA-----PHGNSASAEGRGPSDI---AASART 395
Dd          432 VGNITASATLVNTATTTTPSFYSFTVTVEITMBPSEODEARTDDNNNGPTPVVDWETTNTVT 491
QY          396 AAEGETLESEPAAOVTEVTATSGLVSWGQP 426
Dd          492 SLTPQSTRSTEKTFTIIPVTINDISGI----PG 518

RESULT 12
US-09-903-603A-292
Sequence 292, Application US/09903603A
Patent No. 676795
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
```

APPLICANT: Ferriara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavyn, Ivar J.  
APPLICANT: Mathier, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OR INVENTION: Acids Encoding the Same  
FILE REFERENCE: GNE.1618P2C12  
CURRENT APPLICATION NUMBER: US/09/903,603A  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 292  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-903-603A-292

```

Query Match      12.2%; Score 407.5; DB: 4; Length 640;
Beet Local Similarity 26.8%; Pred. No. 3,78-25;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

Qy          4 P L I L I L A ----- S G --- A A C P L P C V C O N L S S L T C A H R L L V P P V D K R Y E L R   54
              |||::|||
Db          25 P L I V L A L L O L L V A G L V R A Q T C S V C S C S N - O F S R V I C R K R L R E P D G I S T N T R I L N   82

```



```

OY      55 LADFFIALCPPEPPNMTGTVDLTYSRNATITIGARAFGDESLSLSDGNRLVELGTG 11
Db      83 LHEHQOIIKYNSEFKHRLHEITIQLSRNHIRTIEIGAFNGLANITLTPFNRLTTTPNG 1424
OY      115 SLRGPVNLQHLISGN-----QLGRIAAGFDDFIESHD 1494
Db      143 AFVYLSKLKELMRNNPIESIPYAFNRIPSLRJLDGLKRLSLTSSGAEG-LSNDRY 201
OY      150 LDISYNNLRQVPAAGAMPALHTLNDLNLDALPQAFAGLGOLSTL----- 1989
Db      202 LNTLMCMLEIRP--NLPLKLDELDLSGNHLSAIRPSSFQGLMLQLTMMIOSQIOVIE 2583
OY      199 -----DLTSNRLATLAPDLPLFSRGRDAESPAPVLVSFGSNPLHNCCELLML 2454
Db      260 RNAFDNLQSLVEINLANNL-TLPRHDLFTPLHIER-----IHLHNPNCDDTLML 3122
OY      246 R---RLARPDDLTCA--SPGLAGRYFWAVBESEFCEPPLIARHTORLMLVEGORAT 2999
Db      313 SMWIKMABPNTACCRKCNTPPNLKKRYIGEDQYFTCYAPVIEPAPDLNTEGMAAB 3722
OY      300 LRCRALGDPAPTHAWGVDPDDLGVGNS---BARAFPNCTLRIGVAGDAGGYCIATNP 3566
Db      373 LKGR-A-STLSTLSISMTEGNYWTHGAYKRLAIVLSDGLTLFNVTYVQDTEMYTCWNSNS 4311
OY      357 AGEATARVELRVIAL-----PHGNSSAEGRPGPSDI-----AAGART 3955
Db      432 VGNITASATLNTAAATTTTPSYFSTVYVEIMEPSQDEARITDNNVGPFPVVDMEETNVTT 4911
OY      396 AAGEGTLESEPAVQVTEVATNSGLVSNQPG 426
Db      492 SLTPQSTRSTKETTTPVTDINSGL---PG 518

```

```

RESULT 13
US-09-520-781-12
; Sequence 12, Application US/09520781
; Patent No. 6689866
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 No. 668986661 Polynucleotides
; CURRENT APPLICATION NUMBER: US/09/520,781
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: USSN 60/123,667
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 590
; TYPE: prt
; ORGANISM: Homo sapiens
; US-09-520-781-12

```

Query Match	11.6%	Score 386;	DB 4;	length 590;
Best Local Similarity	24.8%	Pred. No. 2e-23;		
Matches	160;	Conservative	77;	Mismatches 247; Indels 160; Gaps 21

  

QY	5	LILLILLASGAA-----CPILPCVCONLESSTLCNAHGLLFVPPNVDRRTVELRLDNFI	60
DB	30	ILCAAIATAAASGPNOCSPVCSGCSN--QFSKYVCTTRGRLSRFPQGISPNSTRYLINIMENNI	87
QY	61	QALGPPDPFNMNGIVDLTLSRNALITRRGARFGDLESKRSIHLIDGNRLVLELGTGSGLRGVY	120
DB	88	QMIQADTRRHLLHLEVLDTQKRSIRQLVEGAFNGLASLSTLEFDNNLTIVIPSGAFEEYLS	147
QY	121	NLOHLLTSGN-----QGLRIAPGAFDDF-----	143
DB	148	KIRELMLRNRPPIESIPSYAFNRVPSLMRLDGLKXLEYISSEGAFBSFLFKLYKINMGCN	207
QY	144	-----LSELELDLSTYNN--LRQVPWAGIGAMPALHTLMLDLALPCGAFQ	191
DB	208	IKDMENLPLVLGTELESNGHFPETIRGSHGSLSLKRLKMYNNSQVSILIR---NAPFG	264

```

QY 192 IGGSTRDITNSRLATLADDPFSGRAKASPAUL-----YLSFGNLIHONCELMYRR 24
Db 265 LASHVELNLAHNNUSL-PHDLT-----PLRYLVEHLHHNNWNOCDCLMTAW 313
QY 248 LAF---PDDLETCA---SPPGLAGRYFAVAFEGEFSCEPPLIARTORTLWLEGORATLR 300
Db 314 WIREYIPTNSTOGRCHAPMHRGRYLYVEVQASQCCAPFIMAPRIMISEGRMAELK 377
QY 302 CRLGSPALPTMAWGEDDRLVGNSS---PAPAFNGTLEIGVTAGAGGYTCIATYPAG 356
Db 374 CRT--PPMSSVKMLFPNGTSLVSHAHRBRISVLNDGTINFSHVLLSDTGVTTCGWTVAG 433
QY 359 EATARVELVNLPHGNGNSAEGRGPSPDIAAASRTAAEGTLSESPANQVTEVATS 418
Db 432 NSNASKYL-----NOSTALNTSNYSFF---TGTBE-----TTEI--- 466
QY 419 GLVSWGGRPADPVMWFQIYNSSEDETLIYRIVPASSHNFLLKHLVPGADYDCLALS 478
Db 465 -----SPDTRTKKKVPVPTSTGY-----Q 484
QY 479 PAAGBEDLTATLLGCAHSTLPASPLCHALQAHV--LGGTLVAVG---GVLVALL 531
Db 485 PATYTTSTVLITTEVPKQAVAPADYTDIKMOTSLDEWAKTKTILIGCFVATLLAAML 544
QY 532 VFVALLVR--GRGANGRLPLKSHVQSCNNGESPSPKHNPR 573
Db 545 IYFYKLRKHQORSTYLAFTLERGHNDKNTPOOKNPOOKOPR 588

```

```

RESULT 14
US-10-140-002-68
; Sequence 68, Application US/10140002
; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 68
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-140-002-68

```

Query Match	11.6%;	Score 385.5;	DB 4;	Length 745;
Best Local Similarity	22.4%;	Pred. No. 3e-23;		
Matches 153;	Conservative 70;	Mismatches 214;	Indels 245;	Gaps 20
QY	6	LLILLISGAACPLPCVCCNLSLSLSTLCARHGLLEFPVPVNDRTTVELRLADNFIQALGP	65	
Db	9	LVALLIGVASSCEPCACVDVKYAHQFADCKYKELRREPE-----GL	49	

```

QY      66 PDERMNTGAVDITLTSNAITTRIGAAFAFGJESIRSLHLDGNRIVELGSGLRGCVNLQHL 125
Db      50 P-----ANVTTLSTLSANKITVLRKGFADVYQVTSLW----- 81
QY      126 IISGNOLGIAPGAFDDPLESLEDIDLSYNNIRQVPMAGICAMPALHTLINDNLIDALP 185
Db      82 -LHNHVRIVTEPGAL-AVLSQIKNDLSHNFISSFPWDELDNLISLQIKNNHRISLSP 139
QY      186 PGAFAPLOGLSRLDLSNRLATLAPDPLPSRGRDAASPAFVLVSFGSNPLHNCCELLMT 245
Db      140 RDLALGALPDLRISIRINNRLFTLAEFT-----DALSLASHLOHTNPHCCGGLVML 199
QY      246 R-----RLARPDDLETCASPRGLAGRYFN-----AVPEGEFSCPEPLIARHTQRL 290
Db      193 QAWAASRVSLEPPDSI-ACASBPALQGVYRYRLPALPCAPPSVHLISAEPLLEAPGT-- 249
QY      291 WVEEGQALTRCAGLADDPAPTMW-----VGPDDRLVGN----- 324
Db      250 -LRAGLAFVLHCLADGHTPTRLQWOLQIPGTVILEPVLVSGEDDVGAAEBGEGGDGL 306
QY      325 -----SSRARAFNGTLEIGVAGAGAGYTCIATNPAGEATARV 364
Db      309 LFTQAGTPTPAPAWPAPRATPRFLALANGSLVLPFLLSAKRGVYTCAHNELGNSTSI 366
QY      365 ELRVLALPHGGSNSAAGCARPGSDIAASKRTAAEBGT----- 404
Db      369 RVAVAAATGPBKIAFGAGGEDDQAPTSEKSTAKRGNISVLPKPEGKIKQGLAKVSL 422
QY      403 --LESEP----- 407
Db      429 GERETPEEDTSGEBAEDQILADPAEORCGNDPSRYVSHAFNQSALPKPHVELGV 486
QY      408 -AVQVTE---VTATSGLVSKGPG-----RPADPVM-----FOIQYNSSD 444
Db      489 IALDVEREARVOLTEPLAARMGPGGAGABRPGRRPLRLYLIPAGGAAYOMSRVEE 546
QY      445 ETILYRVAPASSHHFLKLHPGADYDCLLALSPAGSDTLATR-----LLGCAH 490
Db      549 GUNNY-----MFRGIRPETNYSVCLALAGEBACHQVYVFSTKKEPLSLVIVANSV 596
RESULT 15
US-09-907-794A-73
/ Sequence 73, Application US/09907794A
/ Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Aethkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Macher, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel

```

```

1  APPLICANT: Williams, P. Mickey
2  APPLICANT: Wood, William, I.
3  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
4  TITLE OR INVENTION: Acids Encoding the Same
5  FILE REFERENCE: 10466-14
6  CURRENT APPLICATION NUMBER: US/09/907,794A
7  CURRENT FILING DATE: 2001-07-17
8  PRIOR APPLICATION NUMBER: PCT/US00/04414
9  PRIOR FILING DATE: 2000-02-22
10 PRIOR APPLICATION NUMBER: US 60/143,048
11 PRIOR FILING DATE: 1999-07-07
12 PRIOR APPLICATION NUMBER: US 60/145,698
13 PRIOR FILING DATE: 1999-07-26
14 PRIOR APPLICATION NUMBER: US 60/146,222
15 PRIOR FILING DATE: 1999-07-28
16 PRIOR APPLICATION NUMBER: PCT/US99/20594
17 PRIOR FILING DATE: 1999-09-08
18 PRIOR APPLICATION NUMBER: PCT/US99/20944
19 PRIOR FILING DATE: 1999-09-13
20 PRIOR APPLICATION NUMBER: PCT/US99/21090
21 PRIOR FILING DATE: 1999-09-15
22 PRIOR APPLICATION NUMBER: PCT/US99/21547
23 PRIOR FILING DATE: 1999-09-15
24 PRIOR APPLICATION NUMBER: PCT/US99/23089
25 PRIOR FILING DATE: 1999-10-05
26 PRIOR APPLICATION NUMBER: PCT/US99/28214
27 PRIOR FILING DATE: 1999-11-29
28 PRIOR APPLICATION NUMBER: PCT/US99/28313
29 PRIOR FILING DATE: 1999-11-30
30 PRIOR APPLICATION NUMBER: PCT/US99/28564
31 PRIOR FILING DATE: 1999-12-02
32 PRIOR APPLICATION NUMBER: PCT/US99/28565
33 PRIOR FILING DATE: 1999-12-02
34 PRIOR APPLICATION NUMBER: PCT/US99/30095
35 PRIOR FILING DATE: 1999-12-16
36 PRIOR APPLICATION NUMBER: PCT/US99/30911
37 PRIOR FILING DATE: 1999-12-20
38 PRIOR APPLICATION NUMBER: PCT/US99/30599
39 PRIOR FILING DATE: 1999-12-20
40 PRIOR APPLICATION NUMBER: PCT/US00/00219
41 PRIOR FILING DATE: 2000-01-05
42 NUMBER OF SEQ ID NOS: 423
43 SEQ ID NO 73
44 LENGTH: 620
45 TYPE: PRT
46 ORGANISM: Homo sapiens
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
```

```
Db      323 LAVEPYAFRGLNYLRVLNVSQQLTLEESVPHSVG-NLE-----TLIDSNPLACDC 375
Qy      241 ELIML-RRLARPD---DLFTCASPPGLAGRYFWAVPE---GEFSCEPLIA-RHTORLW 291
Db      376 RLIMVFRRRRLNFNROQPTCATPEFVQKEFDPPDVLBNYFTCRARIRDRKAQOVF 435
Qy      292 VLEGQATLRCRALGDPAPTMHWVGDDRLVGNSSRAR--APPNGTLEIGVTGAGDAGY 349
Db      436 VDEGHTVQFYCRADGDPPEPALIWLSPRKHLVSASKNGRLLTFPPDGLLEVRYAQVDNGTY 495
Qy      350 TCIATNPAGPATRARELRVIAL---PHGNSSAE--GGRPPSDIAASARTAAEGEGL 403
Db      496 LCIAANAGNDSMPAHLHVRSYSPDWPHQNKTFAFISNQPGEGE-ANSTR-----ATV 548
Qy      404 ESEPAVOVTEVTATSGIVSW 423
Db      549 PPFEDIKTULIATTWGFIISF 568
```

Search completed: November 17, 2004, 22:41:41  
Job time : 43 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using SW model

Run on: November 17, 2004, 22:33:09 / Search time 42 Seconds  
(without alignments)  
1454.706 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336

Sequence: 1 MAPPLDLLILASGAACP...LGAGCRGVGSABRLSESVV 635

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:  
2: p1r2:  
3: p1r3:  
4: p1r4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372.5	11.2	421	2	hypothetical prote
2	343.5	10.3	707	2	neuronal leucine-r
3	308	9.2	1091	2	glial cell membran
4	292.5	8.8	789	2	hypothetical prote
5	292.5	8.8	1355	2	hypothetical prote
6	291.5	8.7	361	2	chondroectherin pre
7	287.5	8.6	1535	2	peroxidase - fru
8	279	8.4	1531	2	slit-1 protein hom
9	266.5	8.0	605	2	insulin-like growt
10	265.5	8.0	605	2	insulin-like growt
11	263.5	7.9	622	2	insulin-like growt
12	257.5	7.7	560	2	synleutin - human
13	253.5	7.6	1469	2	platelet membrane
14	253.5	7.6	1480	2	slit protein 1 pre
15	251.5	7.5	603	2	insulin-like growt
16	246	7.4	1523	2	MEG3 protein - ra
17	242	7.3	1025	2	secreted leucine-r
18	239.5	7.2	626	1	platelet glycoprot
19	236.5	7.1	603	2	insulin-like growt
20	221	6.6	907	2	orphan G protein-c
21	213.5	6.4	382	2	proline-arginine
22	213	6.4	420	2	oncogene tropob
23	213	6.4	662	2	garp precursor - h
24	212.5	6.4	575	2	hypothetical prote
25	211.5	6.3	1265	1	neural cell adhesi
26	210.5	6.3	680	2	hypothetical prote
27	207.5	6.2	312	1	leucine-rich alpha
28	207	6.2	375	2	fibromodulin precu
29	206.5	6.2	368	1	biglycan precursor

30	205.5	6.2	1328	2	T23007	hypothetical prote
31	204.5	6.1	369	2	S32793	biglycan precursor
32	203.5	6.1	369	2	S32559	biglycan precursor
33	203.5	6.1	369	2	S20811	proteoglycan I - m
34	203.5	6.1	458	2	T19941	hypothetical prote
35	202.5	6.1	907	2	TG0193	G protein-coupled
36	201	6.0	376	2	S55272	fibromodulin precu
37	198.5	6.0	536	2	A34901	lysine carboxypept
38	194.5	5.8	1097	2	A23943	Toll protein precu
39	192.5	5.8	1256	2	S60461	gene flightless-I
40	191.5	5.7	342	2	A46743	lumican precursor
41	191.5	5.7	738	2	T19938	hypothetical prote
42	191	5.7	1120	2	B86479	hypothetical prote
43	189.5	5.7	380	2	S71876	fibromodulin - ch1
44	187	5.6	4302	2	A38971	polymeric kidney
45	186	5.6	338	2	S52284	lumicon, secretory

## ALIGNMENTS

### RESULT 1

T46266 hypothetical protein DKFp761A179.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C/Accession: T46266

R:Blum, H.; Baurerachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23034

A:Accession: T46266

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <RAA>

A/Cross-references: UNIPROT:Q9NT99; EMBL:AI137451

A/Experimental source: adult amygdala; clone DKFp761A179

A/Genetics:

A/Note: DKFp761A179.1

Query Match 11.2%; Score 372.5; DB 2; Length 421;  
Best Local Similarity 29.0%; Pred. No. 7.7e-17;  
Matches 115; Conservative 41; Mismatches 121; Indels 119; Gaps 14;

QY	67	DFRMTGLVDLTLSNRATIRIGARAFGDLSESLDGNRLVELGTSLRGVNTQHLI 126	
DB	6	DIPNLTALV-----RLEBELSGNRDLIRPGSGQLSRKLW 44	
QY	127	LSGNQLGRIAPGAFDFLESLDLDLSYNNLRQVPWAGIGAMPALHTLNDNLIDALPP 186	
DB	45	LMHAQVATERRAFDD-LKSLLELNLSHNNLMSLP-----HDL----- 81	
QY	187	GAFPAQLGSLRLDLSNRATIRIGARAFGDLSESLDGNRLVELGTSLRGVNTQHLI 246	
DB	82	-FTPLRLERVHL-----NHNPMHCNDVLWLS 108	
QY	247	---RLARPDLLETC---SPGLAGRYFAVGEFSCPEPLIARHTQLRLVLEGORATL 300	
DB	109	KWLKXVPSNTCCARCAPLAKRGYIGELDSHTCAPIYVERPTDLNTEGMAEL 168	
QY	301	KCPALGDPAPYTHMWGPPDRLVGNSS---RARAFTYGLIEIGVTAGAGGYTCIATNPA 357	
DB	169	KCRT-GTSMTSVNWLPNGTLMTHGSYVRISVLHDGTLNFTNVTVQDTGQYTCWVTNSA 227	
QY	358	GEATARVELRYVA---LPHGGSASACGRGPGSDIAASARTAAEGSG-----TL 403	
DB	228	GNTTASATINVAVDPAAGTSGGGGSGSGGSGG-----GSGGYTFTTVEITL 280	
QY	404	ESEPAVQVTEV-----TATSGIWSGPGRPAD 430	
DB	281	ETQPGREALQPRGTEKEPPPTTDGV--WGGRPGD 314	

RESULT 2

JC7763  
neutroal leucine-rich repeat protein-3 - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C/Accession: JC7763  
R/Fukunishi, K.; Matsuo, Y.; Kitahara, C.; Kuchino, Y.; Tsuda, H.  
Biochem. Biophys. Res. Commun. 287, 257-263, 2001  
A/Title: Rat neutroal leucine-rich repeat protein-3: Cloning and regulation of the gene  
A/Reference number: JC7763; PMID:11549284  
A/Contents: Fibrosarcoma cells  
A/Accession: JC7763  
A/Molecule type: mRNA  
A/Residues: 1-707 <FUK>  
A/Cross-references: UNIPROT:Q9RSY6; GB:AF291437  
C/Comment: This protein, a new member of the neutroal leucine-rich repeat protein family  
in protein-protein interaction and functions as a cell adhesion molecule or soluble ligand  
C/Genetics:  
A/Gene: nltr-3  
C/Keywords: cell adhesion

Query Match 10.3%; Score 343.5; DB 2; Length 707;  
Best Local Similarity 24.5%; Pred. No. 1.1e-14;  
Matches 150; Conservative 77; Mismatches 183; Indels 203; Gaps 29;  
QY 17 CPLEPCVCO-----NLSESLSTL-CAHRLGLFVPPNVDRRTVELRLADNFIOAL-GPP 66  
Db CPQLCTCEIRPWPFRPSRYWEASIVDQNDGLNFPARLADDTQILLQTNINLARIHST 88  
QY 67 DFR-NMTGLVDLTLSRNAITRIGARAFGDESLSLHLDGNRLVELGTSLRGPNLQHL 125  
Db DFPVNLGTG--LDLSQNNLSVTNINQKMSQLSVYLEENKLTLPKCYGLSHNLQEL 145  
QY 126 ILISGNQGRAPAF----- 140  
Db 146 YVNNHLTSLAIPGAFVGLHNLRLHLNLSNRLQMINSKWFEALPYLETILMDGNPILRIKD 205  
QY 141 -----DDFLESLEDD-LSTY--NNLRQVPMW----- 163  
Db 206 MNQPLKLRSLVLAGINLVEPPDALVGLLENLESTIFYNRLNKVQVALQKAVNLKFL 265  
QY 164 -----GICAMPALHTLINDHNLIDALP----- 185  
Db 266 DLNKNPINRIRGDFSNMLHKLKELGINNMPDL--VSIDSIAVDNLPILRKITEATNNRSL 323  
QY 186 ---PGAFOIGQSLRLDTLSNRLATLAPDPLFSKRGAEASPAVLVSFGNPLHNCCEL 242  
Db 324 YIHFNAPFLPKESLMLNSNALSYHGTI-----ESLPLKELISHSNPLRCDCVI 376  
QY 243 LMLR-----RLARPDLLETGASPPGLAGR-----YFMAVPEGFSCPEPLIARHT--QR 289  
Db 377 RWINMNTKNTIRFMPDPL-FCVDPPEFGQNVQVHRDMWE--IC-LPLIAPESPSTI 431  
QY 290 LMTLEGORATLRCALDDPAFTMHWVGPD--DRIVGNSSRRAPF--PNGTLEI-GVTGAGD 345  
Db 432 LDVEADSVSLHCGATAEPEPEIYWIITPSSKRLPLNTLREKFYVHSGTLDINGIT-PKE 490  
QY 346 AGGTCTATPAGGATARVELRVLALPHGNSAEGSRPGPSDIASAARTAASEGTTLES 405  
Db 491 GGLYTCTAINLVG-----ADKSTIMYGVGFV-----PQD-----NNGSLN- 526  
QY 406 EPAVQVTEVATSGLVSMGPGRPADPVMFQIQYNS--SEDETLIYRI-VPASSHFL 461  
Db 527 ---IKIRDIRANSVLVSM-----KANSKILKSSVKWTATVKTEDSQAAASRIPSDVKYNNL 580  
QY 462 KHLVPGADYDLCL 474  
Db 581 THLKPSLEYKICI 593

RESULT 3  
A58532  
glial cell membrane glycoprotein LIG-1 precursor - mouse  
C/Species: Mus musculus (house mouse)

C/Date: 11-Apr-1997 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: A58532  
R/Suzuki, Y.; Sato, N.; Tohyama, M.; Kanaka, A.; Takagi, T.  
J. Biol. Chem. 271, 22522-22527, 1996  
A/Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in  
A/Reference number: A58532; MIMD:6394333; PMID:8798419  
A/Accession: A58532  
A/Status: preliminary; translated from GB/EMBL/DBD  
A/Molecule type: mRNA  
A/Residues: 1-1091 <SUZ>  
A/Cross-references: UNIPROT:P70193; GB:D78572; NID:G1545806; PIDN:BA11416.1; PID:G15458  
F/36-61/Domain: proteoglycan amino-terminal homology <PAH>  
F/71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
F/95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
F/118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
F/142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
F/166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
F/191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
F/214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
F/238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
F/262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
F/286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
F/310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
F/334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
F/358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
F/385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
F/409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
F/440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 9.2%; Score 308; DB 2; Length 1091;  
Best Local Similarity 23.9%; Pred. No. 3.9e-12;  
Matches 167; Conservative 63; Mismatches 282; Indels 188; Gaps 23;  
QY 9 ILASGAACPCLPCVCONLSLSLSTLCAHRLGL-----FVPPNVDRRTVELRLADNFIO 61  
Db 179 ILRSGA-----FDGLSRSLTLRLSKRRIQLPYKAKLP---RLQOLDNRNRIR 226  
QY 62 ALGPPFRNMTGLVDLTLSRNAITRIGARAFGDESLSLHLDGNRLVELGTSLRGPNV 121  
Db 227 LIEGLTFQGLDSLEVLRLQNNISRLTDGAFWGLSKNHLVLENSLYEVSNGSLYGLTA 266  
QY 122 LQHLILSGNQLGIAPGAFDDFLESLEDDLSYNNLRQVEMAGIGAMPALHTLINDHNL 181  
Db 287 LHQLHLNNNSISRIQDQW-SFOQKLEHLTISFNMLTRLDSESLAELSSILRLSHNAI 345  
QY 182 DALPFGAFOIGQSLRLDTLSNRLA----- 206  
Db 346 SHIABGAFKGLKSLRVLDLDHNEISGTIEDTSGAFTGLDNLISKLTLFGNKIKSVAKKAFS 405  
QY 207 -----TLAPPLFSRGDAEASPAFL-VLSFGNPLHNCCELWL-----RLLRPD 252  
Db 406 GLESLHNLNGEANAISVQDPAFKKMKLKEYLISSESFLCDQKLKPLPWLGMGRMLQAF 465  
QY 253 DLETGASPPGLAGRYFMAVPEGFSCF--PPLIARHTORLWLEGGORATLRCALGDA 309  
Db 466 VTATCAHPELSLKQGISFVSVPDPSFVCDPDKPOIITQPTTMAVVGKIDIFTGSAASSS 525  
QY 310 PTMHWGPDDRIV-----GNSSRAFAFPNGTLE-----IGVTGADAGGYCTIATNPA 357  
Db 526 SPMTFAKKDNEVLNADMENFAHVRAQDGEVWEYTTILHLRHVTPEHGRYQCLITNHF 585  
QY 358 GEA---TARVELRVL---ALPHGNSAEGRGPGPSDIASAARTAASEGTTLESBPAYQ 410  
Db 586 GSTYSHKARLTVAVLSFTKIPH-----DIARIGTTARLECAATGHPNPQ 631  
QY 411 VTEVATSGLVSMGPGRPADPVMFQIQNSSEDETLIYRIVPASSHFLFKLIVPGADY 470  
Db 632 -----IAW-----QKDGQTD-----PFAERBRM--HYMPDDDV 658  
QY 471 DLCLALSPAAGSDLTATRLGC---AHFTSLPASPLCHALQAVHLGGTLTVAVGVVL 526  
Db 659 FTTTVKIDMDGVYSGTAQNSAGSVSANATLTVALETPSLAVPLEDRV-----VTGE-- 710

QY 527 VAALLVFTVALVVRGAGNGRL-----PLKLSHVQSGTNG----- 562  
 DB 711 -----TVAFOCKATGSPRPRITWLGKGRPLSLTERHHPTPGNQLLVQNVWIDAGRY 763  
 QY 563 -----GPSPTPKAHPPRPPRPORSCSLDGLGACGYA 557  
 DB 764 TCMSNPLGTERAHSSQLSLTPTP--GCRKDGTTVGIFTTA 801

## RESULT 4

T28714

hypothetical protein T21D12.9a - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C/Accession: T28714

R/Moesner, J.

submitted to the EMBL Data Library, August 1997

A/Description: The sequence of *C. elegans* cosmid T21D12.

A/Reference number: Z20514

A/Accession: T28714

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-789 &lt;WOB&gt;

A/Cross-references: EMBL:AF016687; PIDN:AA048096.1; GSPDB:GN00022; CESP:T21D12.9a

A/Experimental source: strain Bristol N2; clone T21D12

C/Genetics:

A/Map position: 4

A/Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

## Query Match

Best Local Similarity 8.8%; Score 292.5; DB 2; Length 789;  
 Matches 136; Conservative 59; Mismatches 225; Indels 133; Gaps 18;

QY 51 VELRLADNFIQALGPDPFRNMTGVLDTLSRNAITRIGARAFDLSLSLHLDGRLYE 110  
 DB 106 VTILKLARNHITTLNQPSFSLRKLRLSLDTRNNIRVRFLAFNQLPSLQNVSLARDVYR 165  
 QY 111 LCTGSLRGVNIQHLILSGNQLGRIPAGAFDDFLEJEDLDLSYNNLRQVPMAGAMP 170  
 DB 166 LDDGMFYACEGLKHLNLSTNRVOAVTEGMFG-LTSLVLDLSYNOIQSFHSSWHTPK 224  
 QY 171 LHTLNDHNLIDALPPGAFQGLS-----RDLTSGNRLA 206  
 DB 225 LKWLSTHSNRIOQLPSGSPFRLRLQLELISANSIDSLHKFALVGMSLHKLDLSNTLA 284  
 QY 207 TLAPDPL-----FSRG-----RDAASAPPLVLTSSGNP--LH----- 237  
 DB 285 VCEVDGAVLYNTSMPLRLSLRFTNNQRLVIRPKAFERFPALBELDLTDNPITATHEAPE 344  
 QY 238 -----CNCCLMLR-----RLARPDDLETCAAPGLAGRYEFAVPEGEF 276  
 DB 345 PLEIKRLVNNSSIIICDCQISWLASWIRYKLKDKSIIKACSYPPPLADLYVAITATNL 404  
 QY 277 SC--EPP--LIARHTQRLWVLEGQATLRCRALGDPATMHW-----VGPDDRVLVG 323  
 DB 405 TCHNDSPRAKIYRQPEVSTLIGEKARFTCNVYGASPLSIEMRVENGOPRVLVQDSATF 464  
 QY 324 NSRRARAFNGTLEIGTVAG-----DAGCTTCIATNPAG--EATARVRLALAP 372  
 DB 465 LSTNRTAVVNGTFDEBELAAELLNDVAMTNDSEYQCVARRNFSGDFSTHYVLQYQAP 524  
 QY 373 HGNSSAEGGRGPSDIAASARTAAEGEGLTSESEPAVOVTEVATSGLVSWGGRPADVY 432  
 DB 525 -----KTYTPEDMPLVLVGQAKLCAATGTPRP--EIK 556  
 QY 433 WMF-QIQNSSSEDETLIYRIVPASSHHFLKLHVPADYDLCIALSPAAGSDLTATRL 491  
 DB 557 WAFEQIIPFAAEARRLY--VTPNDDHIYIMN--VTKEGQAVYCHATNVAAGTOQASA-NL 611  
 QY 492 LGCASFSTLPASP 504  
 DB 612 IVFENFFHYPESP 624

## RESULT 5

T28715

hypothetical protein T21D12.9b - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C/Accession: T28715

R/Moesner, J.

submitted to the EMBL Data Library, August 1997

A/Description: The sequence of *C. elegans* cosmid T21D12.

A/Reference number: Z20514

A/Accession: T28715

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1355 &lt;WOB&gt;

A/Cross-references: EMBL:AF016687; PIDN:AA048095.1; GSPDB:GN00022; CESP:T21D12.9b

A/Experimental source: strain Bristol N2; clone T21D12

C/Genetics:

A/Map position: 4

A/Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2; 786/3; 8

## Query Match

Best Local Similarity 8.8%; Score 292.5; DB 2; Length 1355;  
 Matches 136; Conservative 59; Mismatches 225; Indels 133; Gaps 18;

QY 51 VELRLADNFIQALGPDPFRNMTGVLDTLSRNAITRIGARAFDLSLSLHLDGRLYE 110  
 DB 106 VTILKLARNHITTLNQPSFSLRKLRLSLDTRNNIRVRFLAFNQLPSLQNVSLARDVYR 165  
 QY 111 LCTGSLRGVNIQHLILSGNQLGRIPAGAFDDFLEJEDLDLSYNNLRQVPMAGAMP 170  
 DB 166 LDDGMFYACEGLKHLNLSTNRVOAVTEGMFG-LTSLVLDLSYNOIQSFHSSWHTPK 224  
 QY 171 LHTLNDHNLIDALPPGAFQGLS-----RDLTSGNRLA 206  
 DB 225 LKWLSTHSNRIOQLPSGSPFRLRLQLELISANSIDSLHKFALVGMSLHKLDLSNTLA 284  
 QY 207 TLAPDPL-----FSRG-----RDAASAPPLVLTSSGNP--LH----- 237  
 DB 285 VCEVDGAVLYNTSMPLRLSLRFTNNQRLVIRPKAFERFPALBELDLTDNPITATHEAPE 344  
 QY 238 -----CNCCLMLR-----RLARPDDLETCAAPGLAGRYEFAVPEGEF 276  
 DB 345 PLEIKRLVNNSSIIICDCQISWLASWIRYKLKDKSIIKACSYPPPLADLYVAITATNL 404  
 QY 277 SC--EPP--LIARHTQRLWVLEGQATLRCRALGDPATMHW-----VGPDDRVLVG 323  
 DB 405 TCHNDSPRAKIYRQPEVSTLIGEKARFTCNVYGASPLSIEMRVENGOPRVLVQDSATF 464  
 QY 324 NSRRARAFNGTLEIGTVAG-----DAGCTTCIATNPAG--EATARVRLALAP 372  
 DB 465 LSTNRTAVVNGTFDEBELAAELLNDVAMTNDSEYQCVARRNFSGDFSTHYVLQYQAP 524  
 QY 373 HGNSSAEGGRGPSDIAASARTAAEGEGLTSESEPAVOVTEVATSGLVSWGGRPADVY 432  
 DB 525 -----KTYTPEDMPLVLVGQAKLCAATGTPRP--EIK 556  
 QY 433 WMF-QIQNSSSEDETLIYRIVPASSHHFLKLHVPADYDLCIALSPAAGSDLTATRL 491  
 DB 557 WAFEQIIPFAAEARRLY--VTPNDDHIYIMN--VTKEGQAVYCHATNVAAGTOQASA-NL 611  
 QY 492 LGCASFSTLPASP 504  
 DB 612 IVFENFFHYPESP 624

## RESULT 6

A53860

chondroaderin precursor - bovine

N/Alternate names: 38k leucine-rich protein

C/Species: *Bos primigenius taurus* (cattle)





```

Db      371 DLPRGVFGGLVYTLQILLNANKINCIRPDFAEQDLNLSLSIVDNKIQSLAKGTSTLSRA 430
Qy      226 PLVLSFSGNPINHCNCELLMLRLARPDDLET-----CASPGLAGRYFWMVPEGEFSC 278
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      431 IQLTHLAQNPFICDCKMLKMLADFLKINPIETTTARCAISRRLANKIKGQIKSKKFC 487

RESULT 9
JCS239
insulin-like growth factor acid-labile chain - baboon
C/Spectrum: Papio sp. (baboon)
C/Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C/Accession: JCS239
R/Delhamy, P.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
A/Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like
A/Reference number: JCS239; M0ID:97040714; PMID:8886027
A/Contents: liver
A/Accession: JCS239
A/Molecule type: mRNA
A/Residues: 1-605 <DEL>
C/Comment: This factor is structurally related to proinsulin and have insulin-like meta

Query Match      8.0%; Score 266.5; DB 2; Length 605;
Best Local Similarity 30.9%; Pred. No. 9.3e-10;
Matches 76; Conservative 26; Mismatches 93; Indels 51; Gaps 3;

Qy      16 ACPLPCVCQNLSE--SLSTLCAHRGLLFVPVNVDRRTVELRLADNFIQALGPPDFRNMNG 73
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      40 ACATACACSYDDVENELSFVCSRNILTRLPDGI PGGTQALMDNLSSTIPPAFRLSS 99
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      74 LVDLTISRNAITRIGARFEGDLESRLSLHDGRLVELGTG----- 114
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      100 LAFINLQGGQLSGLEFQALLGLENLCHLBERNQSLASVGTATYTPALALLGNNRSL 159
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      115 -----SLRGPVNLQHILISGNQLGRIAPGAPDFLE 145
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      160 RLEDGLFEGILMDINLGNWSLAVLPDAAFRGLGGLRELIVLAGNRLAVLQPALFSGLA 219
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      146 SLEDLDSYNNLRQVFWAGAMPALHTLNDNLIDALDPGAFAGLGSLRLDLSNRL 205
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      220 -LRELDLSRNALRAIKANVPAQLPRLQKYLDRNLIAA VPGAFLGKALKRLWLDLSHNRV 278
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      206 ATLAPD 211
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      279 AGLLLED 284

RESULT 10
A41915
insulin-like growth factor-binding complex acid-labile chain precursor - human
N/Alternate names: Acid-labile subunit (ALS)
C/Spectrum: Homo sapiens (nan)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: A41915
R/Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.
Mol. Endocrinol. 6, 870-876, 1992
A/Title: Structure and functional expression of the acid-labile subunit of the insulin-like
A/Reference number: A41915; M0ID:92357025; PMID:1379671
A/Accession: A41915
A/Status: preliminary
A/Molecule type: mRNA; protein
A/Residues: 1-605 <LEO>
A/Cross-references: UNIPROT:P35858; GB:M86826; NID:g184807; PIND:AAA36047.1; PID:g184808
A/Experimental source: liver
A/Note: sequence extracted from NCBI backbone (NCBI:P110171)
F/15-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
F/15-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
F/123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F/147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
F/171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
F/195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
F/219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

```

```

F/243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>
F/267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>
F/291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F/315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F/339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F/363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F/387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F/411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F/435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F/459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
F/483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
F/507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>

Query Match      8.0%; Score 265.5; DB 2; Length 605;
Best Local Similarity 32.5%; Pred. No. 1.1e-09;
Matches 80; Conservative 26; Mismatches 89; Indels 51; Gaps 5;

Qy      16 ACPLPCVC--QNLSBSLSTLCAHRGLLFVPVNVDRRTVELRLADNFIQALGPPDFRNMN 72
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      40 ACPAACVCSYDDADDELVSFCSRNILTRLPDGI PGGTQALMDNLSVPPAFAQNLSS 99
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      73 -----GLVDLTISRNAITRIGARFEGDLESRLSLHDGRLV 109
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      100 LQFLINLQGGQLSGLEPQALLGLENLCHLBERNQSLALGTFATTPALASLGLSNNRSL 159
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      110 EL-----GTGSL-----RGPVNLQHILISGNQLGRIAPGAPDFLE 145
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      160 RLEDGLFEGILMDINLGNWSLAVLPDAAFRGLGSLRELIVLAGNRLAVLQPALFSGLA 219
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      146 SLEDLDSYNNLRQVFWAGAMPALHTLNDNLIDALDPGAFAGLGSLRLDLSNRL 205
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      220 -LRELDLSRNALRAIKANVPAQLPRLQKYLDRNLIAA VPGAFLGKALKRLWLDLSHNRV 278
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      206 ATLAPD 211
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      279 AGLLLED 284

RESULT 11
JC7973
synleucin - human
C/Spectrum: Homo sapiens (nan)
C/Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 15-Sep-2003
C/Accession: JC7973
R/Wang, W.; Yang, Y.; Li, L.; Shi, Y.
Biochem. Biophys. Res. Commun. 305, 981-988, 2003
A/Title: Synleucin, a novel leucine-rich repeat protein that increases the intensity o
A/Reference number: JC7973; PMID:12767927
A/Accession: JC7973
A/Molecule type: mRNA
A/Residues: 1-622 <WNA>
A/Cross-references: GB:AY280614
C/Comment: This protein that is a single span transmembrane leucine-rich repeat protei
e intensity of pleiotropic cytokine responses as an adhesion protein or a receptor.
C/Genetics:
A/Map position: 5q12.1
C/Keywords: cytokine; leucine-rich repeat; synleucin; transmembrane protein

Query Match      7.9%; Score 263.5; DB 2; Length 622;
Best Local Similarity 26.7%; Pred. No. 1.5e-09;
Matches 101; Conservative 35; Mismatches 121; Indels 121; Gaps 9;

Qy      6 LILLASGAACPPLPCVCONLSBSLSTLCAHRGLLFVPVNVDRRT----- 50
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      21 LILLHKEITLGC--SSVCG-LCTGQINCRNLGLSIPKPFESTVFLVLTGNNTIYNE 77
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      51 -----VELR-----LADNFIQALGPPDFRMTGIVDL 77
          |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      78 SELTGHSIALVLYLDNSNLTLYPKAFVQLRHLFLFLANNFIRKLDPGIFKGLMLRN 137
          |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      78 TISRNAITRIGARFEGDLESRLSLHDGRLVELGTGSLRGPVNLQHILISGNQLGRIP 137
          |||  |||  |||  |||  |||  |||  |||  |||  |||

```

Db 138 YLVQNVQVSPVPRGVFNLDVSVQVYLNLRNLTVLGSCTFVGMVATRLIDLSNNNIIRISE 197  
 Cy 138 GAPDPLESEEDLDLSTNNLRQVPWAGIGAMPALHTNLNDNLIDALPGAFAGQLSLR 197  
 Db 198 SGF-QHLENIACLTGLGNNILTKVPSNAFEVYKSLRRSLSHNPTEAIQPFKGLANLEY 256  
 Cy 198 LDLTNNRLATLAPD-----PLFSRGRDAE----- 221  
 Db 257 LLLKNSLRVNTRCGFGSINNKLKLLISHNDLENLSDTSLLKNLYLKIDRRIRITSID 316  
 Cy 222 -----ASPAPLVISF-----SGNDLHONCELLWRLRLAP 251  
 Db 317 NDFPENNGASLKIINLTFNNLTLALHPVLKPLSLIHLQANSNMPCNCKLLGLRDWLAS 376  
 Cy 252 DDLETF---CASPRLAGR 266  
 Db 377 SATLNTIYQNPSPMRGR 394

## RESULT 12

A60164  
 Platelet membrane glycoprotein V precursor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 12-Jan-1993 #sequence, revision 24-Feb-1994 #text, change 09-Jul-2004  
 C/Accession: A48030; A60164; A35483; C35483; A60432; A47507; S34329  
 R/Lanza, F.; Morales, M.; de la Salle, C.; Cazenave, J.P.; Clemetson, K.J.; Shimomura, T.  
 J. Biol. Chem. 268, 20801-20807, 1993  
 A/Title: Cloning and characterization of the gene encoding the human platelet glycoprotein V  
 A/Reference number: A48030; MUID:94012616; PMID:8407908  
 A/Accession: A48030  
 A/Molecule type: DNA  
 A/Residues: 1-560 <LA2>  
 A/Cross-references: UNIPROT:P40197; EMBL:Z23091; NID:G312501; PIDN:CAA80637.1; PID:G3125  
 R/Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama, T.  
 Blood 75, 2349-2356, 1990  
 A/Title: Rapid purification and characterization of human platelet glycoprotein V: the a  
 A/Reference number: A60164; MUID:90275263; PMID:2350580  
 A/Accession: A60164  
 A/Molecule type: protein  
 A/Residues: 365-384, 'X', 386-390, 'X', 392-395, 'X', 397, 188-208, 'I', 210, 227-50, 'X', 52-53, 174-  
 'X', 108, 'T', 61-72, 'TK', 75-77, 'V', 56-57, 'G', 479-487, 'X', 489-498, 'X', 500, 'X', 502-503, 'X',  
 R/Roth, G.J.; Church, T.A.; McNeill, B.A.; Williams, S.A.  
 Biochem. Biophys. Res. Commun. 170, 153-161, 1990  
 A/Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related to a  
 A/Reference number: A35483; MUID:90321220; PMID:2372284  
 A/Accession: A35483  
 A/Molecule type: protein  
 A/Residues: 145-166, 'I', 168-169, 'X', 171-172 <ROT>  
 A/Note: this proteolytic fragment was designated peptide M392  
 A/Accession: B35483  
 A/Molecule type: protein  
 A/Residues: 121-129, 'W', 131-135, 466-468, 'X', 470 <RO2>  
 A/Note: this material was designated peptide M393 but may contain two peptides  
 A/Accession: C35483  
 A/Molecule type: protein  
 A/Residues: 252-266, 'H', 268-272, 'X', 274-279, 'I', 281-284, 'I', 286 <RO3>  
 A/Note: this proteolytic fragment was designated peptide M401  
 R/Zafar, R.S.; Walz, D.A.  
 Thromb. Res. 53, 31-44, 1989  
 A/Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive gl  
 A/Reference number: A60432; MUID:89162331; PMID:2922700  
 A/Accession: A60432  
 A/Molecule type: protein  
 A/Residues: 477-478, 'FX', 481-485, 'E', 487, 'V', 489-492, 'NQ', 495, 'E', 497-498 <ZAP>  
 R/Hickey, M.J.; Hagen, F.S.; Yagil, M.; Roth, G.J.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993  
 A/Title: Human platelet glycoprotein V: characterization of the polypeptide and the rela  
 A/Reference number: A47507; MUID:93391348; PMID:7690959  
 A/Accession: A47507  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A/Molecule type: mRNA  
 A/Residues: 1-560 <RES>  
 A/Cross-references: GB:L11238; NID:g388759; PIDN:AAA03069.1; PID:g388760

C/Comment: This platelet membrane protein is a substrate for thrombin.  
 C/Comment: The amino end of the intact protein is blocked.  
 C/Comment: This protein is absent in Bernard-Soulier syndrome.  
 C/Genetics:  
 A/Gene: GDB:GPS  
 A/Cross-references: GDB:230236; OMIM:173511  
 A/Map position: 5pter-5qter  
 C/Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane prot

Query Match 7.7%; Score 257.5; DB 2; Length 560;  
 Best Local Similarity 25.7%; Pred. No. 3.3e-09;  
 Matches 144; Conservative 51; Mismatches 190; Indels 175; Gaps 22;

QY 6 LLLILASGAACPLPCYC-QNISESISTCARGLLFVPPVNDR----- 48  
 Db 10 VIGLRAQPPPCPCPKCFRDAACSGDVARISALGPTNTLTHLLIPMGKGVLOSQS 69  
 QY 49 ---RTV---ELRLADNFIOALGPPDFRMTGLVDLTSRNALITRGARAFGDLESRLHL 103  
 Db 70 FSGMTVLQRLMIDSHLSAVAPGTFSDLTGLKTRLSRNKITHLPGALLKRVLLLEQL 129  
 QY 104 DGNRLVELGTGSLRGPNVLOHLISGNOLGRIAPGAFDPLESLEDLDSYNNLRQVMA 163  
 Db 130 DHVALGIDQNMFOKLVNLQELALNONQDLFLPASLFTN-LENLKLLDLSGNMLTHLPKG 188  
 QY 164 GIGAMPALHTL-----NLDNILDAL-----PGARFOLGQLSRD 199  
 Db 189 LIGAQAKLERLLHSNRLVSLDGLNSLIGALTELOFHRNHIRSIAPGARDRLPNSSLT 248  
 QY 200 LTSNRLATLAPDPLFSRGRDAEASPAVPLVFSGNPLHNCNEV-----LML- 245  
 Db 249 LSNNHIAFL-PSALFLSHNL-----TLTLLENPL--AELPGVLFEGMGGLQELMIN 298  
 QY 246 -----RLARPDDLETCASPPGLAGRYFAVPEGES--CEPLIARHTQRLW 291  
 Db 299 RTQLRTLPAAAFNNLSRLRYLGVTLSP-----RLSALPGARFOLGELQVLAHLSNGLT 352  
 QY 292 VLEGQRATLRCRALGPPAPFMHVVGPDDRLVGNSSSRAEPNG----- 354  
 Db 353 AL-----PDGLRLGILKRVQSLRRNRILRALPRLFNLSLESVQLDHNQ 398  
 QY 335 ---TLEIGVTGA-----GDAGGYTCIATNP--AG 358  
 Db 399 LELTPGVGFALRLTEVLLGHNSWRCDSGLFGLMLRQHLGVG-----BEPPRCAG 453  
 QY 359 EATPARELVALLPHGNSSEAGCGPPSPDIAASAPTAAGSGTL---ESEPAVQTEVT 415  
 Db 454 PG-AHAGLPLMALP-CGDACPCPPRGPPPRPADSSSEAPVHPALAPNSSEPPWMAQPV 511  
 QY 416 ATSGLVSWGPGRPADPVWMP 435  
 Db 512 T-----GKGQDHSPPWGF 524

## RESULT 13

B3665  
 slit protein 2 precursor - fruit fly (Drosophila melanogaster)  
 C/Species: Drosophila melanogaster  
 C/Date: 30-Apr-1991 #sequence, revision 30-Apr-1991 #text, change 02-Aug-2002  
 R/Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.  
 Genes Dev. 4, 2169-2187, 1990  
 A/Title: slit: an extracellular protein necessary for development of midline glia and co  
 A/Reference number: A36665; MUID:91099665; PMID:2176656  
 A/Accession: B3665  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1469 <ROT>  
 A/Cross-references: GB:X53959  
 C/Genetics:  
 A/Gene: FlyBase:sl  
 A/Cross-references: FlyBase:FBgn0003425  
 C/Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

F,66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
 F,101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
 F,125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
 F,149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
 F,173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
 F,197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
 F,228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
 F,288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
 F,323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
 F,347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
 F,371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
 F,395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
 F,419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
 F,450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
 F,512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
 F,547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
 F,572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
 F,596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
 F,620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
 F,651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
 F,708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
 F,743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F,767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
 F,846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
 F,1028-1061/Domain: EGF homology <EGF>  
 F,1068-1099/Domain: EGF homology <EGF2>  
 F,1115-1148/Domain: EGF homology <EGF1>

Query Match 7.6%; Score 253.5; DB 2; Length 1469;  
 Best Local Similarity 20.8%; Pred. No. 1,9e-08;  
 Matches 104; Conservative 47; Mismatches 153; Indels 197; Gaps 11;

15 AACPLEVCVCONSESISTLCAHRLGLFVP---PVDRTTVE----- 52  
 71 ARCPKRCVCTGLNVD---CSHRGLTSVPRKISADVERLELQNNLTVIYETDFQRLTXL 126  
 QY 53 --LRLADNFIQALGPPDFRNMGTGLVDLTLSRAITRIGARAFGLDESLSLHDGRUYE 110  
 DB 127 RMLQLTNDQIHTIERNSPFDIVSLERLDISNNVITTVGRVRFKGAQSLSLQDNNQITC 186  
 QY 111 LGTGSJRGVPVNOHLLISGNQ-----LGR----- 134  
 DB 187 LDEHAFKGLVELEILTINNNTSLPHNIFGGLRRLALRLSDNPPACDCHLSWLSRFLR 246  
 QY 135 ----- 134  
 DB 247 SATRLAPYRCQSPQSKQGNVADLHDFEKCSCGLTEHAPMEGAGNCEPHRCQADGIV 306  
 QY 135 -----IAPGAFDDPLESIEDLISYNNLRQVPMAG 164  
 DB 307 DCKEKSLTSVPVTLPDQTTDVRLQNFITELPKFSFSF--RRLRRIDLNNNISRLAHDA 365  
 QY 165 IGAMPALHTLNDHNLIDLPFGAFQOLGSLDITSNRATLADPLFSKGRDA---- 220  
 DB 366 LSGLKQLTLLVIVGNKIKDLPSCGVFGKLSLRLLLNANETISCTIRDAF---RDLSHS 421  
 QY 221 -----EASPAFLVLSFGNPLHNCCELTMLRRLARPDDLET-----CAS 259  
 DB 422 LLSLYDNNIQLSLANGTFDMKSMKVIHLAKNPICCNIRMLADYHKKPIETISGARCS 481  
 QY 260 PGLAGRYFWAVEPGEFCEPPLIARHTQRLWLVEGQATLRALGADPAFTMHWGPPD 319  
 DB 482 PKMHRRIEISLEEFKCS-----WGELRMKLSGECRMDSDCPAMCHCGTIV 530  
 QY 320 RLVGNSR--AAAFPGTLEI 338  
 DB 531 DCTGRRLKEIPRDIPLHTTEL 551

RESULT 14  
 A3665  
 silt protein 1 precursor - fruit fly (*Drosophila melanogaster*)  
 C/species: *Drosophila melanogaster*

C/Date: 30-Apr-1991 #sequence revision 30-Apr-1991 #text\_change 02-Aug-2002  
 C/Accession: A3665; A31640; S13523  
 R/Author: J.M.; Jacobs, J.R.; Goodman, C.S.; Aravanis-Tsakonas, S.  
 Genes Dev. 4, 2169-2187, 1990  
 A/Title: silt: an extracellular protein necessary for development of midline glia and C  
 A/Reference number: A3665; MUID:9109665; PMID:2176636  
 A/Accession: A3665  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1480 <ROR>  
 A/Cross-references: GB:X53959; NID:98614; PIDN:CAA37910.1; PID:98615  
 R/Author: J.M.; Hartley, D.A.; Walthers, Z.; Aravanis-Tsakonas, S.  
 Cell 55, 1047-1059, 1988  
 A/Title: silt: An EGF-homologous locus of *D. melanogaster* involved in the development  
 A/Reference number: A31640; MUID:8907553; PMID:314436  
 A/Accession: A31640  
 A/Molecule type: DNA  
 A/Residues: 881-1182, 'G', 1185-1404, 'GT', 1463-1464, 'YHA', <ROR>  
 A/Cross-references: GB:M23543; NID:9340939; PID:9514357  
 C/Genetics:  
 A/Genes: FlyBase:silt  
 A/Cross-references: FlyBase:FBgn0003425  
 A/Accession: 1351/3  
 C/Suprafamily: fruit fly silt protein; EGF homology; leucine-rich alpha-2-glycoprotein  
 C/Keywords: alternative splicing; growth factor  
 F,66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
 F,101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
 F,125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
 F,149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
 F,173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
 F,197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
 F,228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
 F,288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
 F,323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
 F,347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
 F,371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
 F,395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
 F,419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
 F,450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
 F,512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
 F,547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
 F,572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
 F,596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
 F,620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
 F,651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
 F,708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
 F,743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F,767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
 F,791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>  
 F,815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>  
 F,846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
 F,1028-1061/Domain: EGF homology <EGF>  
 F,1068-1099/Domain: EGF homology <EGF2>  
 F,1115-1148/Domain: EGF homology <EGF1>

Query Match 7.6%; Score 253.5; DB 2; Length 1480;  
 Best Local Similarity 20.8%; Pred. No. 2e-08;  
 Matches 104; Conservative 47; Mismatches 153; Indels 197; Gaps 11;

15 AACPLEVCVCONSESISTLCAHRLGLFVP---PVDRTTVE----- 52  
 71 ARCPKRCVCTGLNVD---CSHRGLTSVPRKISADVERLELQNNLTVIYETDFQRLTXL 126  
 QY 53 --LRLADNFIQALGPPDFRNMGTGLVDLTLSRAITRIGARAFGLDESLSLHDGRUYE 110  
 DB 127 RMLQLTNDQIHTIERNSPFDIVSLERLDISNNVITTVGRVRFKGAQSLSLQDNNQITC 186  
 QY 111 LGTGSJRGVPVNOHLLISGNQ-----LGR----- 134  
 DB 187 LDEHAFKGLVELEILTINNNTSLPHNIFGGLRRLALRLSDNPPACDCHLSWLSRFLR 246  
 QY 135 ----- 134

```

Db      247 SATRIAPYTRCQSPQLKGNVADLHDPEKCSGLTEHAPMEGAENSCPHPCRCADGIV 306
QY      135 -----IAPGAFDDFLESLEDDLSTYNNLRQVPWAG 164
Db      307 DCEKSLTSVPVTLPPDITVRLQNFITEIPKSFSSF-RRLRRIDLNNNISRIADA 365
QY      165 IGAMPALHTNLNDHNLIDALPPGAFQOLGQLSRLDLSNRLATLAPDPLFSRGRDA----- 220
Db      366 LSGLKQLTTLVLGNKIKDIPSGVFKGLGSLRLILLNANETISCIKDAF----RDLHSL 421
QY      221 -----EASPAPIVLSFGSNPLHNCCELLMRLRLARPDDLET---CAS 259
Db      422 LLSLYDNNISQIANGTFDAKMSKTVHLAKNPICDNLRLADYLHKNPLETSGARCES 481
QY      260 PPLAGGYFNAVPGESCEPPLIARHTQRLWVLEGGRAFLRCALGDPAPTMHWGPPD 319
Db      482 PKMHRRIRESLREKFKCS-----WGELRMKLSGECRMDSDCPAMCHCEGTV 530
QY      320 RLWGNSSR--ARAFNGTLEI 338
Db      531 DCTGRRLKEIPRODPLHTTEL 551

```

## RESULT 15

```

JC1282
Insulin-like growth factor-binding protein acid labile chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: JC1282
R:DaI, J.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 188, 304-309, 1992
A:Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac
A:Reference number: JC1282; MUID:3038676; PMID:1384485
A:Accession: JC1282
A:Molecule type: mRNA
A:Residues: 1-603 <DAI>
A:Cross-references: UNIPROT:P35859; GB:S46785; NID:G258002; PIDN:AA823770.2; PID:G570593
A:Experimental source: liver
A:Note: the authors translated the codon AAG for residue 63 as Arg, AAA for residue 205
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-603/Product: insulin-like growth factor binding protein, acid labile chain #status
F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

```

```

Query Match      7.5%; Score 251.5; DB 2; Length 603;
Best Local Similarity 28.4%; Pred. No. 8.8e-09;
Matches 79; Conservative 38; Mismatches 90; Indels 71; Gaps 6;

```

```

QY      4 PLILLALLA-----SGAA-----CPLPCVC--QNLSESLTCAHGLLF 41
Db      8 PALVVLAFWVVALGPRCHLQGTDPGASADAEGPQCFVACTCSHDDYTDELGVFCSSKVLTH 67
QY      42 VPPVNDKRYTELRADNFICAGLPPDRNMTGLVDLTLSRNATIRIGARAFGDLESIRSL 101
Db      68 LPDDIIVSTRALWDGNNLSISPAAFOQLSSLDFTNLQGSWLRLSLEPQALGLQNLVYL 127
QY      102 HLDGNRLVEIGTGLRGPVNLQHLILSGNQLRIAPGAFDDFLESLEDDLSTYNNLRQVP 161
Db      128 HLEKRNRLNLAVALGFTTTPSLASLSSNLGRLEBGLFOG-LSHWDNLNGWNSLYVLP 186
QY      162 -----MAGIGA-----MPALHT 173
Db      187 DTWFQGLGNLHELVLAKNKLTYLQPALFCGIGELRELDLSRNALRSYKANVFVHLPRLOK 246
QY      174 LNLNDHNLIDALPPGAFQOLGQLSRLDLSNRLATLAPD 211
Db      247 LYLDRNLITAVAGAFILGMKALRWLDLSHNRVAGLMED 284

```

```

Search completed: November 17, 2004, 22:40:56
Job time : 46 secs

```